

Small Ball Estimates for the Fractional Brownian Sheet

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)



vorgelegt dem Rat der Fakultät für
Mathematik und Informatik
der Friedrich-Schiller-Universität Jena

von Dipl.-Math. Thomas Dunker
geboren am 27.1.1971 in Königs Wusterhausen

Contents

Introduction	2
1 Preliminaries	6
1.1 Gaussian Processes and their RKHS	6
General Definitions.	6
The Reproducing Kernel Hilbert Space.	7
The RKHS of \mathbb{B}_d^γ	9
1.2 Schauder Functions	13
Definitions.	13
Some Properties.	14
1.3 Schauder Coefficients for Processes	15
Definitions.	15
Schauder Coefficients.	16
More on Covariance.	18
1.4 Entropy and Approximation Quantities	21
Definitions.	21
Elementary Properties.	21
Mutual Relations.	22
Real and Complex Case.	22
The spaces $H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ and $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$	23
2 Upper Bounds	24
2.1 Key Estimates	24
Hölder-type Norms.	24
Orlicz Norms.	26
2.2 Entropy and Small Ball Results	27
Technical Lemma.	27
Results.	28
2.3 More Entropy and Approximation Results	30
More Function Spaces.	31
Preparations.	32
Results.	34
3 Lower Bounds	36
3.1 A Lower Estimate for the ℓ -Norm	36
3.2 Lower Bounds in the Case $d = 2$	38
Talagrand's Method.	39
Temlyakov's Method.	43
3.3 Lower Bounds for $d > 2$	45
Appendix	47
References	51

Introduction

We consider the fractional Brownian sheet, $\mathbb{B}_d^\gamma := (B_x^\gamma)_{x \in [0,1]^d}$, of order $\gamma \in (0, 2)$ on the d -dimensional unit cube $[0, 1]^d$, $d \geq 2$, i.e. the centered Gaussian random field with covariance

$$\mathbb{E} B_x^\gamma B_y^\gamma = \prod_{j=1}^d \frac{1}{2} (|x_j|^\gamma + |y_j|^\gamma - |x_j - y_j|^\gamma),$$

where $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$. If $\gamma = 1$ then one obtains the usual Brownian sheet. The fractional Brownian sheet is one possible generalization of the fractional Brownian motion on the interval $[0, 1]$. Another generalization is the fractional Levy Brownian field $(W_x^\gamma)_{x \in [0,1]^d}$ which possesses the covariance

$$\mathbb{E} W_x^\gamma W_y^\gamma = \frac{1}{2} (|x|_2^\gamma + |y|_2^\gamma - |x - y|_2^\gamma),$$

where $|\cdot|_2$ denotes the Euclidean norm on \mathbb{R}^d .

We are interested in the asymptotic behavior of the small ball probabilities under Hölder-type and Orlicz norms, more precisely, we wish to find good bounds for

$$-\log \mathbb{P}(\|\mathbb{B}_d^\gamma\| < \varepsilon) \tag{1}$$

as ε tends to zero. For the fractional Levy Brownian field the asymptotic behavior of the small balls under various norms is known. We refer to [Sto96] for results and further references. For the sheet the situation is different.

Concerning applications, the case $\gamma = 1$ – the Brownian sheet – is the most important. Let F be a distribution function of a probability measure on the unit cube $[0, 1]^d$ and consider a sequence X_1, X_2, \dots of independent random variables with distribution function F . Denote by $[0, x]$, $x \in [0, 1]^d$, the set $\prod_{j=1}^d [0, x_j]$. Then the empirical distribution functions F_n , $n = 1, 2, \dots$, are defined as

$$F_n(x) := \frac{1}{n} \sum_{l=1}^n \mathbf{1}_{[0,x]}(X_l)$$

and it is known that the processes $(\sqrt{n}(F_n(x) - F(x)))_{x \in [0,1]^d}$ converge in law (with respect to the sup-norm) to a centered Gaussian field $\mathbb{Y}_F := (Y_x)_{x \in [0,1]^d}$ with covariance

$$\mathbb{E} Y_x Y_y = F([0, x] \cap [0, y]) - F([0, x])F([0, y]).$$

Suppose now that $X_1, \dots, X_n \in [0, 1]^d$ are samples obtained by a sequence of independent experiments. In order to test whether these samples could have the distribution function F one computes the F_n defined above and the Kolmogorov-Smirnov statistic $\sqrt{n}(F_n - F)$. If later is too large one would like to reject the hypothesis that X_1, \dots, X_n have the distribution function F .

Since the distribution of the empirical distribution functions are difficult to determine one uses the process \mathbb{Y}_F as approximation. If one wants to evaluate the power of this test then one is confronted with the problem of estimating

$$\mathbb{P}(\|\sqrt{n}(F_n(x) - F(x))\|_\infty < \varepsilon)$$

when X_1, \dots, X_n have a distribution function $\tilde{F} \neq F$. This leads us to:

$$\mathbb{P}(\|\sqrt{n}(F_n - F)\|_\infty < \varepsilon) \approx \mathbb{P}(\|\mathbb{Y}_{\tilde{F}} - \sqrt{n}(\tilde{F} - F)\|_\infty < \varepsilon)$$

– a small ball problem with a shifted ball. It has been shown that

$$\begin{aligned} \mathbb{P}(\|\tilde{\mathbb{Y}}_{\tilde{F}} - \sqrt{n}(\tilde{F} - F)\|_\infty < \varepsilon/2) &\leq \mathbb{P}(\|\mathbb{Y}_{\tilde{F}} - \sqrt{n}(\tilde{F} - F)\|_\infty < \varepsilon) \\ &\leq \frac{C}{\varepsilon} \mathbb{P}(\|\tilde{\mathbb{Y}}_{\tilde{F}} - \sqrt{n}(\tilde{F} - F)\|_\infty < 2\varepsilon) \end{aligned}$$

where $\tilde{\mathbb{Y}}_{\tilde{F}} := (\tilde{Y}_x)_{x \in [0,1]^d}$ is the centered Gaussian field with covariance

$$\mathbb{E} \tilde{Y}_x \tilde{Y}_y = \tilde{F}([0, x] \cap [0, y]).$$

Finally, consider the case that \tilde{F} is the uniform distribution on the unit cube $[0, 1]^d$. Then one observes that $\tilde{\mathbb{Y}}_{\tilde{F}}$ is nothing but the Brownian sheet \mathbb{B}_d^1 . And if \tilde{F} possesses a density \tilde{f} then we have the representation

$$\tilde{\mathbb{Y}}_{\tilde{F}} = \int \mathbf{1}_{[0, \cdot]}(x) \sqrt{\tilde{f}(x)} dB_x.$$

Thus we see that the Brownian sheet plays an important role for a large class of examples. For more details we refer to [Bas88] where we took this example from.

Another type of question where small ball estimates are needed are Chung-type laws of the iterated logarithm. If we take into account that we deal with self-similar processes, i.e. $(B_x^\gamma)_{x \in [0,1]^d} \stackrel{d}{=} (T^{-d\gamma/2} B_{Tx}^\gamma)_{x \in [0,1]^d}$, the problem can be formulated as follows. Let $M(\cdot, \cdot)$ be a functional on $C([0, \infty)^d) \times (0, \infty)$ such that $M(\cdot, T)$ is a norm on $C([0, T]^d)$ and $M(f, T) = T^\beta M(f(T\cdot), 1)$ for some $\beta \in \mathbb{R}$. We search for a normalization function ψ such that

$$\liminf_{T \rightarrow \infty} \frac{T^{\beta+d\gamma/2} M(\mathbb{B}_d^\gamma, 1)}{\psi(T)}$$

is almost surely in $(0, \infty)$. For example, in [Tal94] M. Talagrand found the asymptotic small ball behavior for the two-dimensional Brownian sheet under the sup-norm and obtained the normalization function

$$\psi(T) = \frac{T(\log \log \log T)^{3/2}}{(\log \log T)^{1/2}}.$$

for the Chung-type law of the iterated logarithm. For more Chung-type results and techniques we refer to [MR95] and [Sto95].

For the fractional Brownian sheet we could not find references in the literature. Thus, the following brief historical review is completely devoted to small ball estimates for the standard Brownian sheet. For $d = 1$ and $\gamma = 1$, where we deal with the classical Wiener process, the exact asymptotic behavior of (1) is known. In this

case, using techniques of differential equations one has several series representations for the probability (1), see e.g. [Fel64], Chapter 10. In particular,

$$-\log \mathbb{P}(\|\mathbb{B}_1^1\|_{C([0,1])} < \varepsilon) \sim \frac{\pi^2}{8\varepsilon^2}.$$

We refer to [BM91] for L_p -extensions of this result.

In the multiparameter case $d > 1$ it is not clear how to use differential equations and a completely different technique is needed. First estimates for (1) in the multiparameter case date back only to 1979 when P. Révész [Rév81] proved for $d = 2$ the existence of constants C_1, C_2 such that

$$\frac{C_1}{\varepsilon^2} |\log \varepsilon| \leq -\log \mathbb{P}(\|\mathbb{B}_2^1\|_{C([0,1]^2)} < \varepsilon) \leq \frac{C_2}{\varepsilon^2} |\log \varepsilon|^5 \quad (2)$$

for small ε . In 1982 E. Csáki [Csá82] found the asymptotic behavior of the small ball probability of \mathbb{B}_d^1 under the $L_2([0,1]^d)$ -norm. He showed that

$$-\log \mathbb{P}(\|\mathbb{B}_d^1\|_{L_2([0,1]^d)} < \varepsilon) \sim \frac{K_d^2}{\varepsilon^2} |\log \varepsilon|^{2d-2}$$

with constant

$$K_d := \frac{2^{d-2}}{\sqrt{2}\pi^{d-1}(d-1)!}.$$

See [Li92] for various non-Brownian multiparameter generalizations of this result. Using the inequality $\|\cdot\|_{L_2([0,1]^d)} \leq \|\cdot\|_{C([0,1]^d)}$ one obtains

$$-\log \mathbb{P}(\|\mathbb{B}_d^1\|_{C([0,1]^d)} < \varepsilon) \geq \frac{C}{\varepsilon^2} |\log \varepsilon|^{2d-2},$$

which improves in the case $d = 2$ the lower estimate in (2). The next result on this problem gave improved upper bounds. In 1986 M. A. Lifshits ([LT86], for $d = 2$) and in 1988 R. F. Bass ([Bas88], for general d) obtained

$$-\log \mathbb{P}(\|\mathbb{B}_d^1\|_{C([0,1]^d)} < \varepsilon) \leq \frac{C}{\varepsilon^2} |\log \varepsilon|^{3d-3}. \quad (3)$$

At that stage, a considerable gap of order $d - 1$ remained between the exponents of the log-terms in lower and upper bounds. In 1994 M. Talagrand [Tal94] succeeded in proving the sharpness of (3) for $d = 2$. Yet, it is not known how to generalize this result to higher dimension $d \geq 3$.

In this work, we will apply a technique which was developed in a recent work [DKLL98] for the case $\gamma = 1$. This method improved the upper estimate (3) to

$$-\log \mathbb{P}(\|\mathbb{B}_d^1\|_{C([0,1]^d)} < \varepsilon) \leq \frac{C}{\varepsilon^2} |\log \varepsilon|^{2d-1}$$

which reduces the gap between the exponents of the log-terms in upper and lower bounds to one, independently of the dimension $d \geq 3$. It appears that this technique

is powerful also in a more general setting. Our method is based on a remarkable paper by J. Kuelbs and W. V. Li [KL93], where they proved that the small ball behavior of a Gaussian process is closely connected with the entropy numbers of an operator mapping from a Hilbert space H into a Banach space which is an isometry between H and the reproducing kernel Hilbert space of the process. A second ingredient is that we can analyze the processes using Schauder function expansions. This is a frequently used tool in the theory of almost surely continuous Gaussian processes in order to investigate their path properties. We refer to [CKR93] and [Sto96].

This work is organized as follows. In Chapter 1 we introduce all the necessary notions and prove some preliminary facts. We start with some basic concepts from the theory of Gaussian processes. Then we discuss connections between classical function spaces and the reproducing kernel Hilbert spaces of the d -dimensional fractional Brownian sheets. Next, we provide all the facts about Schauder functions which we need and we conclude this chapter presenting the necessary facts about entropy and approximation quantities.

Chapter 2 is devoted to upper estimates for (1). First, we are concerned with estimates of Kolmogorov and entropy numbers for certain integral operators. Then the small ball estimates for the processes \mathbb{B}_d^γ are obtained via the relation mentioned above. We close this chapter with a list of entropy results for embeddings.

In Chapter 3 we collect some results which give us lower bounds. First, we show that the key estimate of our method is sharp for the Brownian sheet. Next, we present results which are due to M. Talagrand and V. N. Temlyakov and which prove the sharpness of the upper estimates for the case $d = 2$. We give a new proof for the main inequality of Talagrand's method which appears to be much shorter. Let us remark that both methods use the special situation $d = 2$ and it is not obvious how to treat higher dimensions. Finally, we discuss our results for $d > 2$.

Acknowledgment. I would like to thank very much my academic advisor W. Linde who stimulated and supported this study. For interesting discussion on this subject I owe many thanks to T. Kühn, M. Lifshits and W. V. Li.

Next, I would like to thank my parents who supported my development and helped me to find my way. And, I don't want to forget to mention all my friends who cheered me up when my research gave me a hard time.

Note. After this work had been finished we were informed that E. S. Belinsky had obtained similar result independently. His results contain the first part of our Corollary 2.11. He showed for $p \in (1, \infty)$ and $\alpha = (\alpha_1, \dots, \alpha_d)$ with $\max(1/p, 1/2) < \alpha_1 = \dots = \alpha_\nu < \alpha_{\nu+1} \leq \dots \leq \alpha_d$ that

$$e_k(\iota : H_p^\alpha(\mathbb{T}^d) \rightarrow C(\mathbb{T}^d)) \lesssim k^{-\alpha_1} (1 + \log k)^{\alpha_1(\nu-1)+1/2}$$

(for the explanation of the notation we refer to the next chapter). This and other results can be found in his preprint entitled "Estimates of Entropy Numbers and Gaussian Measures for Classes of Functions with Bounded Mixed Derivative" which he has submitted to the Journal of Approximation Theory.

1 Preliminaries

We start with some remarks on frequently used notions. By C, C_1, \dots we will denote constants which can be different at each occurrence. Let f and g be two non-negative functions which are defined on an infinite sequence or on some interval on the positive half line. We will write $f \lesssim g$ if there exists a constant $C > 0$ such that $f \leq Cg$. Analogously, we will use the notation $f \gtrsim g$ when $g \lesssim f$. If both relations $f \lesssim g$ and $f \gtrsim g$ hold true then we will write $f \approx g$.

1.1 Gaussian Processes and their RKHS

The concepts of Gaussian process and Gaussian field have already appeared in the introduction. Let us use this section to recall some definitions and facts concerning this subject.

General Definitions. Let I be an arbitrary index set. Then the indexed set of random variables, $\mathbb{X} = (X_t)_{t \in I}$ with $X_t : \Omega \rightarrow \mathbb{R}$, is a Gaussian process if for all $n \in \mathbb{N}$ and all $t_1, \dots, t_n \in I$ the finite dimensional distributions $\text{dist}(X_{t_1}, \dots, X_{t_n})$ are Gaussian distributions on \mathbb{R}^n , i.e. for any choice of real numbers a_1, \dots, a_n there exist a $\mu \in \mathbb{R}$ and a $\sigma \geq 0$ such that for all $s \in \mathbb{R}$

$$\mathbb{P}(a_1 X_{t_1} + \dots + a_n X_{t_n} < s) = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^s \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right) dt$$

and we say $a_1 X_{t_1} + \dots + a_n X_{t_n}$ is normally distributed with mean μ and variance σ^2 , short $\mathcal{N}(\mu, \sigma^2)$. Since the n -dimensional Gaussian distributions are uniquely determined by their mean function $m(t) := \mathbb{E} X_t$ and their covariance function $K(s, t) := \mathbb{E}(X_s - m(s))(X_t - m(t))$ it follows that the distribution of a Gaussian process is uniquely determined by these two functions, too. A process is called centered if its mean function is constant zero.

Most of the time the index or parameter set I is some interval on the real line. In our situation I will be the d -dimensional unit cube $[0, 1]^d$. Processes with multidimensional parameter sets are often called random fields.

Now, let us look on the process $\mathbb{X} = (X_t)_{t \in [0, 1]^d}$ in a different way. We fix an $\omega \in \Omega$ and consider the trajectory $\mathbb{X}(\omega) = X(\omega)$ as a function over the cube $[0, 1]^d$. We say that \mathbb{X} has almost surely continuous trajectories if there exists a measurable subset $\Omega' \subset \Omega$ with full measure such that $\mathbb{X}(\omega)$ is continuous for all $\omega \in \Omega'$. Obviously, this property is very fragile. One can change the values of the X_t 's on sets of measure zero such that the process is not anymore almost surely continuous. Yet, in the same way one can “repair” the trajectories without changing the distribution of the process. That's why one usually requires only that the process possesses a version \mathbb{X}' , i.e. $\mathbb{P}(X_t = X'_t) = 1$ for all $t \in [0, 1]^d$, which has almost surely continuous trajectories.

Having a process $\mathbb{X} = (X_t)_{t \in [0, 1]^d}$ with almost surely continuous trajectories one can interpret \mathbb{X} as Gaussian random variable taking values in the Banach space $C([0, 1]^d)$. In general, a random variable $\mathbb{X} : \Omega \rightarrow E$ with values in a separable

Banach space E is Gaussian if for each functional $a \in E'$ the \mathbb{R} -valued random variable $a(\mathbb{X})$ is normally distributed. Again, one knows that the distribution of such a Gaussian variable is uniquely determined by the mean function $m(a) := \mathbb{E} a(\mathbb{X})$, $a \in E'$, and the covariance function $K(a, b) := \mathbb{E} (a(\mathbb{X}) - m(a))(b(\mathbb{X}) - m(b))$ where $a, b \in E'$.

The Reproducing Kernel Hilbert Space. Let us now consider a centered Gaussian random variable \mathbb{X} taking values in a separable Banach space E . The reproducing kernel Hilbert space (short: RKHS) is defined as

$$\mathcal{H}_{\mathbb{X}} := \{x \in E : \exists c > 0 \forall a \in E' a(x)^2 \leq c^2 K(a, a)\}$$

equipped with the norm

$$\|x\|_{\mathbb{X}} := \sup_{\substack{a \in E' \\ K(a, a) \neq 0}} \frac{|a(x)|}{\sqrt{K(a, a)}}.$$

The unit ball $\mathcal{B}_{\mathbb{X}}$ of $\mathcal{H}_{\mathbb{X}}$ is a compact set in E . Conversely, the RKHS determines the Gaussian distribution of \mathbb{X} on E . Obviously, we have

$$K(a, a) = \sup_{x \in \mathcal{B}_{\mathbb{X}}} a(x)^2.$$

As we mentioned in the introduction, the small ball problem for Gaussian random variables is closely linked to the entropy numbers of the compact set $\mathcal{B}_{\mathbb{X}}$. Namely, we can use the following theorem which is due to W. V. Li and W. Linde [LL]. It improves the earlier result [KL93] by J. Kuelbs and W. V. Li.

Theorem 1.1 (W. V. Li, W. Linde) *Let $\alpha > 1/2$ and let $J(t)$ be a slowly varying function for t tending to infinity. We consider the following four statements:*

- (a) *it holds $-\log \mathbb{P}(\|\mathbb{X}\|_E < \varepsilon) \gtrsim \varepsilon^{-2/(2\alpha-1)} J(1/\varepsilon)^{2/(2\alpha-1)}$,*
- (b) *for $\mathcal{B}_{\mathbb{X}} \subset E$ we have $e_k(\mathcal{B}_{\mathbb{X}}) \gtrsim k^{-\alpha} J(k)$,*
- (c) *it is $-\log \mathbb{P}(\|\mathbb{X}\|_E < \varepsilon) \lesssim \varepsilon^{-2/(2\alpha-1)} J(1/\varepsilon)^{2/(2\alpha-1)}$ and*
- (d) *we have $e_k(\mathcal{B}_{\mathbb{X}}) \lesssim k^{-\alpha} J(k)$,*

where e_k denotes the k -th (dyadic) entropy number (for definition see Section 1.4). The statements (a) and (b) are equivalent and (c) implies (d). If, in addition, $J(t)$ is a non-decreasing function and, for all $\rho > 0$, we have $J(t^\rho) \approx J(t)$, as t tends to infinity, then statement (c) follows from (d).

Hence it is necessary to study the structure of the RKHS of our processes in order to be able to apply known entropy results or to derive new ones from the small ball estimates.

Before we proceed with our special processes, let us review some general facts about RKHS's of Gaussian random variables. We denote by $H_{\mathbb{X}} \subset L_2(\Omega, \mathbb{P})$ the Hilbert space

$$H_{\mathbb{X}} := \text{cl}(\{a(\mathbb{X}) : a \in E'\})$$

where the closure is taken with respect to the $L_2(\Omega, \mathbb{P})$ -norm. Note that $H_{\mathbb{X}}$ is separable. Let $T : H_{\mathbb{X}} \rightarrow E$ denote the operator which is defined by the Bochner integral $Tg := \mathbb{E} g \mathbb{X}$ for $g \in H_{\mathbb{X}}$. Then one verifies easily that

$$a(Tg)^2 = (\mathbb{E} ga(\mathbb{X}))^2 \leq E g^2 K(a, a)$$

and, hence, $T(H_{\mathbb{X}}) \subset \mathcal{H}_{\mathbb{X}}$. Conversely, for $x \in \mathcal{H}_{\mathbb{X}}$ the functional $L_x(a(\mathbb{X})) := a(x)$ is bounded on $\{a(\mathbb{X}) : a \in E'\}$. Consequently, it can be extended continuously to $H_{\mathbb{X}}$. Then by Riesz's Theorem there exists an element $g \in H_{\mathbb{X}}$ such that $a(x) = \mathbb{E} ga(\mathbb{X}) = a(\mathbb{E} g \mathbb{X})$ for all $a \in E'$. This implies $x = \mathbb{E} g \mathbb{X} \in T(H_{\mathbb{X}})$. Furthermore, it holds $\langle Ta(\mathbb{X}), Tb(\mathbb{X}) \rangle_{\mathbb{X}} = \mathbb{E} a(\mathbb{X})b(\mathbb{X})$ which shows that T is an isometry between the two Hilbert spaces $H_{\mathbb{X}}$ and $\mathcal{H}_{\mathbb{X}}$.

An immediate consequence of the above considerations is the so called Karhune-Loève expansion. Let f_1, f_2, \dots be an arbitrary complete orthonormal system in $\mathcal{H}_{\mathbb{X}}$ then we can represent the random variable \mathbb{X} as a series

$$\mathbb{X} \stackrel{d}{=} \sum_{n=1}^{\infty} g_n f_n$$

where g_1, g_2, \dots are independent $\mathcal{N}(0, 1)$ -distributed random variables. And, if the operator $\tilde{T} : H \rightarrow \mathcal{H}_{\mathbb{X}}$ is an isometry between some Hilbert space H and the RKHS of \mathbb{X} and h_1, h_2, \dots is some complete orthonormal system in H then we have

$$\mathbb{X} \stackrel{d}{=} \sum_{n=1}^{\infty} g_n \tilde{T} h_n.$$

Consider now the case that $E = C(I)$ where I is a compact subset of \mathbb{R}^d . Obviously, we have $(T\delta_s(\mathbb{X}))(\cdot) = \mathbb{E} \delta_s(\mathbb{X})\delta(\cdot) = K(s, \cdot)$. By the bilinearity of K it follows that $\text{span}\{T\delta_s(\mathbb{X}) : s \in I\} = \text{span}\{K(s, \cdot) : s \in I\}$ where T is the operator defined above. Since the functionals $\text{span}\{\delta_s : s \in I\}$ are dense in $C'(I)$ with respect to the weak topology, we can find for each $a \in C'(I)$ a sequence $\{a_n\} \subset \text{span}\{\delta_s : s \in I\}$ which converges weakly to a . By the Banach-Steinhaus Theorem the norms of the a_n 's are uniformly bounded by some constant. Then the integrability of Gaussian norms and Lebesgue's Dominated Convergence Theorem yield

$$\begin{aligned} \lim_{n \rightarrow \infty} \|Ta(\mathbb{X}) - Ta_n(\mathbb{X})\|_{\mathbb{X}}^2 &= \lim_{n \rightarrow \infty} \mathbb{E} (a(\mathbb{X}) - a_n(\mathbb{X}))^2 \\ &= \mathbb{E} \lim_{n \rightarrow \infty} (a(\mathbb{X}) - a_n(\mathbb{X}))^2 = 0. \end{aligned}$$

Thus, we can conclude $\mathcal{H}_{\mathbb{X}} = \overline{\text{span}}\{K(s, \cdot) : s \in I\}$ where the closure is taken with respect to the norm induced by the scalar product $\langle K(s, \cdot), K(t, \cdot) \rangle_{\mathbb{X}} = K(s, t)$.

The RKHS of \mathbb{B}_d^γ . We consider the functions

$$K_t^\gamma(s) := K^\gamma(s, t) = \frac{1}{2}(|t|^\gamma + |s|^\gamma - |t - s|^\gamma), \quad s \in [0, 1],$$

for $t \in [0, 1]$. Then it follows from the discussion above that the RKHS $\mathcal{H}_{\mathbb{B}_1^\gamma}$ is given by

$$\mathcal{H}_{\mathbb{B}_1^\gamma} := \overline{\text{span}}\{K_t^\gamma : t \in [0, 1]\}, \quad (1.1)$$

where the closure is taken with respect to the norm which is induced by the scalar product $\langle \cdot, \cdot \rangle_{\mathbb{B}_1^\gamma}$, which is defined by $\langle K_s^\gamma, K_t^\gamma \rangle_{\mathbb{B}_1^\gamma} = K^\gamma(s, t)$. For \mathbb{B}_d^γ , one can verify that

$$(\mathcal{H}_{\mathbb{B}_d^\gamma}, \langle \cdot, \cdot \rangle_{\mathbb{B}_d^\gamma}) = (\mathcal{H}_{\mathbb{B}_1^\gamma}, \langle \cdot, \cdot \rangle_{\mathbb{B}_1^\gamma})^{\otimes d}.$$

Thus, it suffices to concentrate on the case $d = 1$. The description (1.1) is not very convenient and we wish to have a characterization in terms of classical function spaces. For this purpose let us recall some concepts from the theory of function spaces.

Some Concepts in Function Spaces. Throughout the rest of this section we will deal mainly with complex function spaces. This is more convenient and the corresponding results for real spaces are obtained by restriction to the real subspaces.

We denote by $S(\mathbb{R})$ the Schwartz space of all rapidly decreasing, infinitely differentiable functions on \mathbb{R} . Its topology is generated by the family of norms

$$\|\varphi\|_{k,l} := \sup_{t \in \mathbb{R}} (1 + |t|)^k \sum_{j=0}^l |D^j \varphi(t)|$$

$k, l = 0, 1, \dots$. Thus, $S(\mathbb{R})$ is a complete locally convex space. Let $S'(\mathbb{R})$ be its topological dual, the space of all tempered distributions. For $\varphi \in S(\mathbb{R})$, the Fourier transform $\mathcal{F}\varphi$ is defined by

$$\mathcal{F}\varphi(t) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(s) \exp(ist) ds \quad t \in \mathbb{R},$$

and by $\mathcal{F}f(\varphi) := f(\mathcal{F}\varphi)$, for all $\varphi \in S(\mathbb{R})$, the Fourier transform is extended to $S'(\mathbb{R})$. Recall that \mathcal{F} is a surjective isomorphism on $S(\mathbb{R})$ and $S'(\mathbb{R})$, respectively, and \mathcal{F} restricted to $L_2(\mathbb{R})$ is a unitary transform.

Let us consider Bessel potential spaces of order $\alpha \geq 0$ in the Hilbert space case. They are defined as

$$H_2^\alpha(\mathbb{R}) := \{f \in S'(\mathbb{R}) : (1 + |\cdot|^2)^{\alpha/2} \mathcal{F}f \in L_2(\mathbb{R})\}$$

equipped with the norm

$$\|f\|_{H_2^\alpha(\mathbb{R})} := \left(\int_{\mathbb{R}} (1 + |t|^2)^\alpha |\mathcal{F}f(t)|^2 dt \right)^{1/2}.$$

These spaces coincide with the spaces $W_2^\alpha(\mathbb{R})$ which are called Slobodeckij spaces or Sobolev spaces if α is an integer. Other equivalent descriptions of these spaces are the Besov spaces $\Lambda_{2,2}^\alpha(\mathbb{R})$ and the spaces $F_{2,2}^\alpha(\mathbb{R})$ and $B_{2,2}^\alpha(\mathbb{R})$. Their definitions can be found e.g. in [Tri83] on the pages 36/37 and 45, respectively.

On the interval $(0, 1)$ one defines $H_2^\alpha((0, 1))$ by restriction. Let $D((0, 1))$ be the space of all infinitely differentiable functions with compact support in $(0, 1)$. Its topology is defined as follows: $\{\varphi_n : n \in \mathbb{N}\} \subset D((0, 1))$ converges to $\varphi \in D((0, 1))$ if there exists a compact subset of $(0, 1)$ such that all supports of the φ_n 's and φ are contained in this set and $D^j \varphi_n$ converges uniformly to $D^j \varphi$ for all $j = 0, 1, \dots$. Now, denote by $D'((0, 1))$ the topological dual of $D((0, 1))$. Then $H_2^\alpha((0, 1))$ is defined as

$$H_2^\alpha((0, 1)) := \{f \in D'((0, 1)) : \exists g \in H_2^\alpha(\mathbb{R}) \text{ } g|_{(0,1)} = f\}$$

and

$$\|f\|_{H_2^\alpha((0,1))} := \inf\{\|g\|_{H_2^\alpha(\mathbb{R})} : g \in H_2^\alpha(\mathbb{R}), g|_{(0,1)} = f\}.$$

Finally, we consider the torus $\mathbb{T} = [0, 1)$ and denote by $D(\mathbb{T})$ the space of all periodic infinitely differentiable functions. Its topology is induced by the family of semi-norms

$$\|\varphi\|_k := \sup_{t \in \mathbb{T}} |D^k \varphi(t)|$$

$k = 0, 1, \dots$. Then, we denote by $D'(\mathbb{T})$ the topological dual space of $D(\mathbb{T})$. The periodic Bessel potential space $H_2^\alpha(\mathbb{T})$ is the space

$$\{f \in D'(\mathbb{T}) : ((1 + |k|^2)^{\alpha/2} \hat{f}(k))_{k \in \mathbb{Z}} \in l_2(\mathbb{Z})\}$$

and

$$\|f\|_{H_2^\alpha(\mathbb{T})} := \left(\sum_{k \in \mathbb{Z}} (1 + |k|^2)^\alpha |\hat{f}(k)|^2 \right)^{1/2}$$

where $\hat{f}(k)$ denotes the k -th Fourier coefficient $f(\exp(-i2\pi k \cdot))$.

Let us close this excursion on function spaces by stating the following useful proposition.

Proposition 1.2 *Assuming $1/2 < \alpha < 3/2$, the completion of $D((0, 1))$ under the norm $\|\cdot\|_{H_2^\alpha(\mathbb{R})}$ coincides with the space*

$$H_{2,0}^\alpha((0, 1)) := \{f \in H_2^\alpha((0, 1)) : f(0) = f(1) = 0\}$$

and it can be interpreted as

$$H_{2,0}^\alpha(\mathbb{T}) := \{f \in H_2^\alpha(\mathbb{T}) : f(0) = 0\}.$$

Let us add some words on this proposition. First of all, the condition $f(0) = 0$ makes sense since the spaces $H_2^\alpha((0, 1))$ embed into the space of continuous functions provided $\alpha > 1/2$. The first statement of Proposition 1.2 can be found in [Tri83], Corollary 3.4.3. For the second part of Proposition 1.2 we give a sketch of the proof in the Appendix. We chose this way since we could not find an easy

reference.

We come back to the spaces $\mathcal{H}_{\mathbb{B}_1^\gamma}$. In the case $\gamma = 1$ one can easily find an isometry between $L_2((0, 1))$ and $\mathcal{H}_{\mathbb{B}_1^1}$. One can verify that the mapping defined by $T : \mathbf{1}_{[0,t]} \mapsto K_t^1$ is such an isometry since $\{\mathbf{1}_{[0,t]} : t \in [0, 1]\}$ is dense in $L_2((0, 1))$. Observe that T is nothing but the integration operator $Tf(t) := \int_0^t f(s) ds$.

For general $\gamma \in (0, 2)$, the situation is less straightforward. We need an integral operator T mapping some $L_2(M, \mu)$ into $C([0, 1])$ with kernel $k(x, s)$, $s \in [0, 1]$ and $x \in M$, such that $\langle k(\cdot, s), k(\cdot, t) \rangle_{L_2(M, \mu)}$ equals the covariance $K(s, t)$ of the process. Then T is defined as $Tf(t) := \int_M f(x)k(x, t) \mu(dx)$. Having this in mind one can check that the fractional Fourier transform $\mathcal{F}_\gamma : L_2(\mathbb{R}) \rightarrow C([0, 1])$, which is defined as

$$\mathcal{F}_\gamma f(t) := \int_{\mathbb{R}} \frac{\exp(its) - 1}{|s|^{(\gamma+1)/2}} f(s) ds$$

(see Singer [Sin94]), is such an example which maps in the following way

$$\frac{1}{c_\gamma} \mathcal{F}_\gamma : \frac{1}{c_\gamma} \frac{\exp(it\cdot) - 1}{|\cdot|^{(\gamma+1)/2}} \mapsto K_t^\gamma.$$

Applying the Fourier transform one obtains the more classical operator

$$\frac{1}{c'_\gamma} T^\gamma : \frac{1}{c'_\gamma} ((\cdot - t)_-^{(\gamma-1)/2} - (\cdot)_-^{(\gamma-1)/2}) \mapsto K_t^\gamma$$

((\cdot) $_-$ can be replaced by $|\cdot|$ which effects only the constant). In the paper of L. Decreusefond and A. S. Üstünel [DÜ] one can also find isometries between $L_2((0, 1))$ and $\mathcal{H}_{\mathbb{B}_1^\gamma}$. Let us work with the fractional Fourier transform. Theorem 3.1 from [Sin94] provides us with the following properties:

(F1) for $f \in L_2(\mathbb{R})$ we have $|\mathcal{F}_\gamma f(t)| \leq c_\gamma^2 |t|^{\gamma/2} \|f\|_{L_2(\mathbb{R})}$,

(F2) let $\varphi \in D(\mathbb{R} \setminus \{0\})$ then the following inversion formula holds

$$\mathcal{F}_\gamma^{-1} \varphi(t) := \frac{1}{\sqrt{2\pi}} |t|^{(\gamma+1)/2} \mathcal{F}^{-1} \varphi(t)$$

where \mathcal{F}^{-1} denotes the usual inverse Fourier transform.

We decompose $\mathcal{H}_{\mathbb{B}_1^\gamma}$ into

$$\mathcal{H}_{\mathbb{B}_1^\gamma, 0} := \{f - f(1)K_1^\gamma : f \in \mathcal{H}_{\mathbb{B}_1^\gamma}\} = \{f \in \mathcal{H}_{\mathbb{B}_1^\gamma} : f(0) = f(1) = 0\}$$

and $\text{span}\{K_1^\gamma\}$ which are mutually orthogonal. Then we can state the following lemma.

Lemma 1.3 *For $\gamma \in (0, 2)$, the spaces $\mathcal{H}_{\mathbb{B}_1^\gamma, 0}$ and $H_{2,0}^{(\gamma+1)/2}((0, 1))$ coincide.*

Proof. Let $L_2(\mathbb{R})^\gamma \subset L_2(\mathbb{R})$ denote the subspace which is isometrically mapped onto $\mathcal{H}_{\mathbb{B}_1^\gamma,0}$ by $c_\gamma^{-1}\mathcal{F}_\gamma$, that is

$$L_2(\mathbb{R})^\gamma := \overline{\text{span}}\left\{\frac{1}{c_\gamma} \frac{\exp(it\cdot) - 1}{|\cdot|^{(\gamma+1)/2}} - K^\gamma(t,1) \frac{1}{c_\gamma} \frac{\exp(i\cdot) - 1}{|\cdot|^{(\gamma+1)/2}} : t \in [0,1]\right\},$$

where the closure is taken with respect to the $L_2(\mathbb{R})$ -norm.

First we show that $D((0,1)) \subset \mathcal{H}_{\mathbb{B}_1^\gamma,0}$. Let $\varphi \in D((0,1))$ and set $f := c_\gamma \mathcal{F}_\gamma^{-1} \varphi$. We decompose f into $f_1 \in L_2(\mathbb{R})^\gamma$ and $f_2 \in (L_2(\mathbb{R})^\gamma)^\perp$. Assume that $c_\gamma^{-1} \mathcal{F}_\gamma f_2(t) \neq 0$ for some $t \in (0,1)$, i.e

$$\frac{1}{c_\gamma} \left\langle \frac{\exp(it\cdot) - 1}{|\cdot|^{(\gamma+1)/2}}, f_2 \right\rangle_{L_2(\mathbb{R})} = K^\gamma(t,1) \frac{1}{c_\gamma} \left\langle \frac{\exp(i\cdot) - 1}{|\cdot|^{(\gamma+1)/2}}, f_2 \right\rangle_{L_2(\mathbb{R})} \neq 0,$$

then it follows that $c_\gamma^{-1} \mathcal{F}_\gamma f_2(1) \neq 0$, too. Yet $c_\gamma^{-1} \mathcal{F}_\gamma f_1(1) = 0$ and this contradicts $\varphi \in D((0,1))$. Consequently, $f = f_1$ which implies that $\varphi \in \mathcal{H}_{\mathbb{B}_1^\gamma,0}$.

Secondly, we observe that $c_\gamma \mathcal{F}_\gamma^{-1}(D((0,1)))$ is dense in $L_2(\mathbb{R})^\gamma$. In order to see this, take a $f \in L_2(\mathbb{R})^\gamma$, with $f \not\equiv 0$, and define the distribution $L_f \in D'((0,1))$

$$L_f(\varphi) := \int_{\mathbb{R}} f(s) |s|^{(\gamma+1)/2} \mathcal{F}^{-1} \varphi(s) ds, \quad \varphi \in D((0,1)),$$

then $f \perp c_\gamma \mathcal{F}_\gamma^{-1}(D((0,1)))$ iff $\text{supp } L_f \subset \{0,1\}$. Hence L_f is a sum of derivatives of the distributions δ_0 and δ_1 . Applying the Fourier transform we deduce that f must be of the form $f(t) = (p_1(t) + \exp(it)p_2(t))/|t|^{(\gamma+1)/2}$ where p_1 and p_2 are polynomials. Since $f \in L_2(\mathbb{R})$ it follows that $\deg p_i < \gamma/2 < 1$ and $p_1(0) = -p_2(0)$ which implies that $f(t) = c(\exp(it) - 1)/|t|^{(\gamma+1)/2} \in (L_2(\mathbb{R})^\gamma)^\perp$ which contradicts the assumption.

Finally, using (F1) and (F2), we see that on $D((0,1))$ we have the estimate

$$\|(1 + |\cdot|^2)^{(\gamma+1)/4} \mathcal{F} \varphi\|_{L_2(\mathbb{R})} \leq \|\varphi\|_{L_2(\mathbb{R})} + \| |\cdot|^{(\gamma+1)/2} \mathcal{F}^{-1} \varphi \|_{L_2(\mathbb{R})} \leq C \|\varphi\|_{\mathcal{H}_{\mathbb{B}_1^\gamma}}.$$

Obviously, the inverse inequality is true, too, and, consequently, both norms are equivalent. Then the assertion is a consequence of Proposition 1.2. \square

Corollary 1.4 *The RKHS $\mathcal{H}_{\mathbb{B}_d^\gamma}$ of the d -dimensional fractional Brownian sheet of order γ , with $\gamma \in (0,2)$, can be interpreted as*

$$(H_{2,0}^\alpha(\mathbb{T}) \oplus \text{span}\{K_1^\gamma\})^{\otimes d},$$

where $\alpha := (\gamma + 1)/2 \in (1/2, 3/2)$. In particular, this contains the space

$$H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) := \left\{ f \in D'(\mathbb{T}^d) : f(x_1, \dots, x_d) = 0, \text{ whenever } \exists j \text{ with } \right. \\ \left. x_j = 0 \text{ and } \sum_{k \in \mathbb{Z}^d} \prod_{j=1}^d (1 + |k_j|^2)^\alpha |\hat{f}(k)|^2 < \infty \right\}$$

– a space with dominating mixed smoothness.

1.2 Schauder Functions

Definitions. First, we consider the system of Schauder functions on the interval $[0, 1]$. It consists of the functions

$$u_{-2,0} := \mathbf{1}_{[0,1]}, \quad u_{-1,0}(s) := s, \quad s \in [0, 1]$$

and functions $u_{m,i}$ which are generated as follows. Let u be the piecewise linear function

$$u(s) := s\mathbf{1}_{(0,1/2)}(s) + (1-s)\mathbf{1}_{[1/2,1)}(s), \quad s \in \mathbb{R}, \quad (1.2)$$

then we define

$$u_{m,i}(s) := 2^{-m/2}u(2^m(s - i2^{-m})), \quad s \in [0, 1], \quad (1.3)$$

where $m = 0, 1, \dots$ and $i = 0, \dots, 2^m - 1$. We remark that the index m corresponds to a dilation while the index i corresponds to a shift. Let us denote by $J_m := \{0, \dots, 2^m - 1\}$, for $m = 0, 1, \dots$, and by $J_m := \{0\}$, for $m = -2, -1$, the sets of all possible shifts in the m -th level. From (1.2) and (1.3), one sees that for fixed m the functions $u_{m,i}$ are supported by dyadic intervals of length $2^{-\max(m,0)}$, i.e. the supports of two different $u_{m,i}$'s are essentially disjoint.

As usual, $C([0, 1])$ denotes the space of all continuous functions on the interval $[0, 1]$ and it is equipped with the supremum norm. One can verify that for any $f \in C([0, 1])$ we have a unique Schauder function expansion

$$f(s) = \sum_{m=-2}^{\infty} \sum_{i \in J_m} f_{m,i} u_{m,i}(s), \quad s \in [0, 1]$$

where the coefficients are $f_{-2,0} = f(0)$, $f_{-1,0} = f(1) - f(0)$ and

$$f_{m,i} = 2^{m/2} \left(2f\left(\frac{2i+1}{2^{m+1}}\right) - f\left(\frac{i}{2^m}\right) - f\left(\frac{i+1}{2^m}\right) \right), \quad (1.4)$$

for $m = 0, 1, \dots$ and $i \in J_m$ (see e.g. [KS89], Chapter VI.). In order to simplify some later formulas we introduce the following functionals and operators. For $0 \leq s < t \leq 1$, we define the functional $\Delta_{s,t}$ on $f \in C([0, 1])$ by $\Delta_{s,t} := f(t) - f(s)$. Furthermore, we will use the abbreviations $\Delta^{-2,0}(f) := f(0)$, $\Delta^{-1,0} := \Delta_{0,1}$ and

$$\Delta^{m,i} := 2^{m/2} \left(\Delta_{\frac{i}{2^m}, \frac{2i+1}{2^{m+1}}} - \Delta_{\frac{2i+1}{2^{m+1}}, \frac{i+1}{2^m}} \right),$$

for $m = 0, 1, \dots$ and $i \in J_m$. Let $S_{m,i} : C([0, 1]) \rightarrow C([0, 1])$ be the one-dimensional operator

$$S_{m,i}f := (\Delta^{m,i} \otimes u_{m,i})f = \Delta^{m,i}(f)u_{m,i}$$

and denote by S_m the operator $\sum_{i \in J_m} S_{m,i}$ which is of rank $2^{\max(0,m)}$. Finally, we set

$$\mathbf{S}_1^n := \sum_{m=-2}^n S_m, \quad n = -2, -1, 0, 1, \dots$$

Observe that $(\mathbf{S}_1^n f)_{n \in \mathbb{N}}$ forms a sequence of piecewise linear approximations of f . One can check that for all points $s \in \{i2^{-(n+1)} : i = 0, \dots, 2^{n+1}\}$ we have $\mathbf{S}_1^n f(s) = f(s)$ and $\mathbf{S}_1^n f$ is linear on the dyadic intervals of length $2^{-(n+1)}$.

Now, let us turn to the multi-dimensional case $C([0, 1]^d)$, $d \geq 2$. We extend the above concepts by taking tensor products. In the sequel, let m denote a multi-index $(m_1, \dots, m_d) \in \{-2, -1, \dots\}^d$ and let $J_m := J_{m_1} \times \dots \times J_{m_d}$. We define

$$|m| := \sum_{j=1}^d \max(0, m_j)$$

and we call $|m|$ the order of the multi-index m . The multidimensional Schauder functions are defined as $u_{m,i} := u_{m_1,i_1} \otimes \dots \otimes u_{m_d,i_d}$, where $m \in \{-2, -1, \dots\}^d$ and $i \in J_m$. Analogously, we write $\Delta_{x,y}$ for the functional $\Delta_{x_1,y_1} \otimes \dots \otimes \Delta_{x_d,y_d}$, where $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ are in $[0, 1]^d$ and $x_j < y_j$ for all $j = 1, \dots, d$. For $m \in \{-2, -1, \dots\}^d$ and $i \in J_m$, we set $\Delta^{m,i} := \Delta^{m_1,i_1} \otimes \dots \otimes \Delta^{m_d,i_d}$ and we define

$$S_{m,i} := \Delta^{m,i} \otimes u_{m,i} = S_{m_1,i_1} \otimes \dots \otimes S_{m_d,i_d}.$$

Summing over all $i \in J_m$ we obtain $S_m := \sum_{i \in J_m} S_{m,i}$ which is an operator of rank $2^{|m|}$. We denote by \mathcal{S}^n , $n = 0, 1, \dots$, the operator $\sum_{|m| \leq n} S_m$. Note that \mathcal{S}^n is not the d -fold tensor product of the operator \mathbf{S}_1^n . This would be the operator

$$\mathbf{S}_d^n := \sum_{m \in \{-2, \dots, n\}^d} S_m = (\mathbf{S}_1^n)^{\otimes d}.$$

Some Properties. Let us collect below some elementary properties of the d -dimensional Schauder functions. For $m \in \{-2, -1, \dots\}^d$, we define

$$\overline{m} = (\overline{m}_1, \dots, \overline{m}_d) := (\max(0, m_1), \dots, \max(0, m_d)).$$

Clearly, it holds $|m| = |\overline{m}|$. A first observation is that

$$\lambda^d(\text{supp } u_{m,i}) = 2^{-|m|} \quad \text{and} \quad \|u_{m,i}\|_{C([0,1]^d)} \leq 2^{-|m|/2}, \quad (1.5)$$

where λ^d denotes the d -dimensional Lebesgue measure. Let us denote by b_m , $m \in \{0, 1, \dots\}^d$, the “brick”

$$b_m := \prod_{j=0}^d [0, 2^{-m_j}). \quad (1.6)$$

We note that the support of the Schauder function $u_{m,i}$ is essentially the brick $(i_1 2^{-\overline{m}_1}, \dots, i_d 2^{-\overline{m}_d}) + b_{\overline{m}}$. For fixed m , the bricks $(i_1 2^{-\overline{m}_1}, \dots, i_d 2^{-\overline{m}_d}) + b_{\overline{m}}$, $i \in J_m$, form a partition of the cube $[0, 1]^d$ which we want to denote by

$$P_{\overline{m}} := \{(i_1 2^{-\overline{m}_1}, \dots, i_d 2^{-\overline{m}_d}) + b_{\overline{m}} : i \in J_m\}. \quad (1.7)$$

Finally, let us compute the rank of the operators $\mathcal{S}^n - \mathcal{S}^{n-1}$. For this purpose, we need to know the cardinality of the set $\{m \in \{-2, -1, \dots\}^d : |m| = n\}$. Obviously, we have

$$\begin{aligned} & \#\{m \in \{-2, -1, \dots\}^d : |m| = n\} \\ &= \sum_{j=1}^d 2^{d-j} \binom{d}{d-j} \#\{m \in \{0, 1, \dots\}^j : |m| = n\} \end{aligned} \quad (1.8)$$

and $\#\{m \in \{0, 1, \dots\}^j : |m| = n\}$ can be computed recursively using the formula

$$\#\{m \in \{0, 1, \dots\}^j : |m| = n\} = \sum_{k=0}^n \#\{m \in \{0, 1, \dots\}^{j-1} : |m| = n - k\}.$$

By induction one can prove that

$$\#\{m \in \{0, 1, \dots\}^j : |m| = n\} = \binom{n+j-1}{j-1}. \quad (1.9)$$

Substituting (1.9) in (1.8) yields

$$\#\{m \in \{-2, -1, \dots\}^d : |m| = n\} = \sum_{j=1}^d 2^{d-j} \binom{d}{d-j} \binom{n+j-1}{j-1} \approx n^{d-1}.$$

Consequently, we obtain

$$\text{rank}(\mathcal{S}^n - \mathcal{S}^{n-1}) = \sum_{|m|=n} \text{rank } S_m \approx n^{d-1} 2^n. \quad (1.10)$$

1.3 Schauder Coefficients for Processes

Definitions. We study Gaussian random fields on the unit cube $[0, 1]^d$ with covariance functions which can be written as tensor product of a covariance function over the interval $[0, 1]$. In the sequel, let $K : [0, 1]^2 \rightarrow \mathbb{R}$ be a covariance function of a centered Gaussian process $\mathbb{G}_1 := (G_t)_{t \in [0, 1]}$. We denote by $\delta_t \in C'([0, 1])$ the functional $\delta_t(f) := f(t)$, which is often called δ -distribution in the point t . It is obvious that K can be regarded as bilinear functional on $\text{span}\{\delta_t : t \in [0, 1]\}$ and we will use the notation

$$K(a, b) := \mathbb{E} a(\mathbb{G}_1) b(\mathbb{G}_1)$$

where $a, b \in \text{span}\{\delta_t : t \in [0, 1]\}$. For our purposes, it will be more convenient to deal with increments of the process \mathbb{G}_1 . That's why we introduce the functional

$$\begin{aligned} \sigma^2(s, t) = \sigma^2(t, s) &:= \mathbb{E} \Delta_{s,t}(\mathbb{G}_1)^2 \\ &= K(s, s) - 2K(s, t) + K(t, t) \end{aligned}$$

for $s, t \in [0, 1]$ with $s \leq t$. Let $s \leq t$ and $s' \leq t'$, then one can check that

$$\mathbb{E} \Delta_{s,t}(\mathbb{G}_1) \Delta_{s',t'}(\mathbb{G}_1) = \frac{1}{2} (\sigma^2(s, t') + \sigma^2(s', t) - \sigma^2(s, s') - \sigma^2(t, t')).$$

Now, we consider the centered Gaussian random field $\mathbb{G}_d := (G_x)_{x \in [0,1]^d}$, $d \geq 2$, which possesses the covariance function $K^{\otimes d}$, i.e.

$$K^{\otimes d} \left(\bigotimes_{j=1}^d a_j, \bigotimes_{j=1}^d b_j \right) := \prod_{j=1}^d K(a_j, b_j),$$

where $a_j, b_j \in \text{span}\{\delta_t : t \in [0, 1]\}$ for all $j = 1, \dots, d$. Note that the tensor product of positive definite functions is again positive definite. The d -dimensional fractional Brownian sheet of order $\gamma \in (0, 2)$ is a special example of such a Gaussian random field.

Schauder Coefficients. Following the ideas from [Sto96] we introduce the following conditions on σ^2 . Let γ be in the interval $(0, 2)$:

- ($\sigma 1$) $\sigma^2(s, t) \leq C_1 |t - s|^\gamma$, for all $s, t \in [0, 1]$ and some constant C_1 ,
- ($\sigma 2$) $\sigma^2(s, t) \geq C_2 |t - s|^\gamma$, for all $s, t \in [0, 1]$ and some constant C_2 ,
- ($\sigma 3$) $\sigma^2(s, t) \leq c[\sigma^2(s, (s+t)/2) + \sigma^2((s+t)/2, t)]$, for all $s, t \in [0, 1]$ and some constant $c \in (0, 2)$.

The first condition ($\sigma 1$) implies that \mathbb{G}_d admits a version which has almost surely continuous trajectories. This is a consequence of the estimate

$$\begin{aligned} & \mathbb{E} (G_x - G_y)^{2N} \\ & \leq (2N - 1)!! \left(\sum_{j=1}^d \left(\mathbb{E} (G_{(y_1, \dots, y_{j-1}, x_j, \dots, x_d)} - G_{(y_1, \dots, y_j, x_{j+1}, \dots, x_d)})^2 \right)^{1/2} \right)^{2N} \\ & \leq (2N - 1)!! C_1^N \|K\|_{C([0,1]^2)}^N \left(\sum_{j=1}^d |x_j - y_j|^{\gamma/2} \right)^{2N} \\ & \leq (2N - 1)!! C_1^N \|K\|_{C([0,1]^2)}^N d^{1-\gamma/4} |x - y|_2^{\gamma N} \end{aligned}$$

and Kolmogorov's continuity theorem (see [LT91], Corollary 11.4). Thus, we are in the situation that we can apply all the concepts developed in Section 1.2 for continuous functions. Then, the coefficients $g_{m,i} := \Delta^{m,i}(\mathbb{G}_d)$ are Gaussian random variables in \mathbb{R} . The next lemma collects all necessary formulas which are needed in order to determine the covariance of the $g_{m,i}$'s.

Lemma 1.5 *For $m, m' \geq 0$, $i \in J_m$ and $i' \in J_{m'}$ we introduce the abbreviations $t := (2i + 1)2^{-(m+1)}$, $t' := (2i' + 1)2^{-(m'+1)}$, $h := 2^{-(m+1)}$ and $h' := 2^{-(m'+1)}$. Then we have*

$$\begin{aligned} & \mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m',i'}(\mathbb{G}_1) = \\ & -2^{(m+m')/2-1} \left[\left(\sigma^2(t-h, t'-h') - 2\sigma^2(t, t'-h') + \sigma^2(t+h, t'-h') \right) \right. \\ & \quad \left. - 2 \left(\sigma^2(t-h, t') - 2\sigma^2(t, t') + \sigma^2(t+h, t') \right) \right. \\ & \quad \left. \left(\sigma^2(t-h, t'+h') - 2\sigma^2(t, t'+h') + \sigma^2(t+h, t'+h') \right) \right] \end{aligned} \quad (1.11)$$

and, in particular,

$$\mathbb{E} \Delta^{m,i}(\mathbb{G}_1)^2 = 2^m (2\sigma^2(t-h, t) + 2\sigma^2(t, t+h) - \sigma^2(t-h, t+h)). \quad (1.12)$$

In the remaining cases, where m equals -1 or -2 , one can compute

$$\begin{aligned} \mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{-1,0}(\mathbb{G}_1) = \\ -2^{m/2-1} [\sigma^2(t-h, 0) - 2\sigma^2(t, 0) + \sigma^2(t+h, 0)) \\ - (\sigma^2(t-h, 1) - 2\sigma^2(t, 1) + \sigma^2(t+h, 1))], \end{aligned} \quad (1.13)$$

$$\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{-2,0}(\mathbb{G}_1) = 2^{m/2} (2K(0, t) - K(0, t-h) - K(0, t+h)),$$

where $m \geq 0$, $i \in J_m$ and t, h as above, and, finally, $\mathbb{E} \Delta^{-1,0}(\mathbb{G}_1)^2 = \sigma^2(0, 1)$, $\mathbb{E} \Delta^{-2,0}(\mathbb{G}_1)^2 = K(0, 0)$ and $\mathbb{E} \Delta^{-1,0}(\mathbb{G}_1) \Delta^{-2,0}(\mathbb{G}_1) = K(0, 1) - K(0, 0)$.

For the Gaussian random field \mathbb{G}_d , $d \geq 2$, the covariances can be calculated by using the formula

$$\begin{aligned} \mathbb{E} \Delta^{m,i}(\mathbb{G}_d) \Delta^{m',i'}(\mathbb{G}_d) &= \prod_{j=1}^d K(\Delta^{m_j, i_j}, \Delta^{m'_j, i'_j}) \\ &= \prod_{j=1}^d \mathbb{E} \Delta^{m_j, i_j}(\mathbb{G}_1) \Delta^{m'_j, i'_j}(\mathbb{G}_1) \end{aligned} \quad (1.14)$$

for all $m, m' \in \{-2, -1, \dots\}^d$, $i \in J_m$ and $i' \in J_{m'}$.

These formulas can be checked easily by elementary calculations. Then substitution in (1.12) yields the following corollary.

Corollary 1.6 *The first condition $(\sigma 1)$ implies that*

$$\mathbb{E} \Delta^{m,i}(\mathbb{G}_d)^2 \leq C 2^{(1-\gamma)|m|}$$

for all $m \in \{-2, -1, \dots\}^d$ and $i \in J_m$. If, in addition, K satisfies the conditions $(\sigma 2)$ and $(\sigma 3)$ then there exists a constant C' such that

$$\mathbb{E} \Delta^{m,i}(\mathbb{G}_d)^2 \geq C' 2^{(1-\gamma)|m|} \quad (1.15)$$

for all $m \in \{0, 1, \dots\}^d$ and $i \in J_m$. If $K(0, 1) - K(0, 0) \neq 0$ then (1.15) holds for all $m \in \{-1, 0, \dots\}^d$ and if also $K(0, 0) \neq 0$ then inequality (1.15) is true for all $m \in \{-2, -1, \dots\}^d$.

Remark. Suppose that $\sigma^2(s, t) = |s - t|$ and $K(0, 0) = 0$, then one can check immediately from formula (1.11)–(1.14) that the coefficients $g_{m,i}$, with $m \in \{-1, 0, \dots\}^d$ and $i \in J_m$, are independent $\mathcal{N}(0, 1)$ -distributed.

Remark. Given a norm the next natural question is: Does the Schauder expansion with the coefficients investigated above converge almost surely? Since the answer will be a simple consequence of our key estimates in Section 2.1 we shall not discuss it here.

More on Covariance. In [Sto96], it was shown for $d = 1$ that under special additional assumptions on σ^2 the covariance matrix of the vector $(g_{m,i})_{i=0}^{2^m-1}$ is in a certain sense diagonal dominant, i.e. the sum of the entries in one row is bounded by the element from the diagonal times a universal constant. Furthermore, one can reduce the vector by taking only every 2^k -th element and obtains diagonal dominance for the covariance of the reduced vector. If, in addition, one has norms which can be estimated from below by some sequence space norm of the Schauder coefficients then one can compute lower bounds for the asymptotic behavior of $-\log \mathbb{P}(\|\mathbb{G}_1\| < \varepsilon)$. This last point is our problem. We did not find good inequalities which estimate our norms under consideration from below by means of Schauder coefficients. Nevertheless, we would like to show below under which condition one has “diagonal dominance” for the vectors $(g_{m,i})_{|m|=n, i \in J_m}$. Maybe, this can be useful for some later work.

Lemma 1.7 *We assume that $(\sigma 1)$ is satisfied for some $\gamma \in (0, 2)$ and suppose that $K(0, 0) = 0$. In addition we require that*

$$(\sigma 4) \quad |D^{(2,0)}\sigma^2(s, t)| \leq C_3|s - t|^{\gamma-2} \text{ and } |D^{(2,1)}\sigma^2(s, t)| \leq C_4|s - t|^{\gamma-3} \text{ hold for all } s, t \in [0, 1] \text{ and some constants } C_3 \text{ and } C_4.$$

Then we have for any $m \in \{-1, 0, \dots\}^d$

$$\sum_{|m'|=|m|} \sum_{i' \in J_{m'}} |\mathbb{E} \Delta^{m,i}(\mathbb{G}_d) \Delta^{m',i'}(\mathbb{G}_d)| \leq C 2^{(1-\gamma)|m|} \quad (1.16)$$

where C is a constant depending only on d and γ .

Proof. The right hand sides of (1.11) and (1.13) consist of summands of the form

$$R(t, h, s) := \sigma^2(t - h, s) - 2\sigma^2(t, s) + \sigma^2(t + h, s),$$

namely, we have for $m, m' \geq 0$

$$\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m',i'}(\mathbb{G}_1) = -2^{(m+m')/2-1} (R(t, h, t' - h') - 2R(t, h, t') + R(t, h, t' + h'))$$

and

$$\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{-1,0}(\mathbb{G}_1) = 2^{m/2-1} (R(t, h, 1) - R(t, h, 0)),$$

where we used the abbreviations t, t', h and h' defined in Lemma 1.5. We distinguish several cases

Case 1 ($m = m'$). For $m = -1$ and $m = 0$ the index set J_m contains only one element and $\mathbb{E} \Delta^{m,0}(\mathbb{G}_1)^2 \leq C$. Now, let $m \geq 1$. For $|t - t'| \leq 2^{-m+1}$ we simply use the Cauchy Schwarz inequality and get

$$|\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m,i'}(\mathbb{G}_1)| \leq C 2^{(1-\gamma)m}$$

by Corollary 1.6. Next, we investigate the situation where $|t - t'| > 2^{-m+1}$. By Taylor's formula we can find $\theta_1, \theta_2 \in [0, 1]$ such that

$$R(t, h, t' + h) - R(t, h, t') = 2^{-3(m+1)} (D^{(2,1)}\sigma^2(t + \theta_1 h, t' + \theta_2 h) + D^{(2,1)}\sigma^2(t - \theta_1 h, t' + \theta_2 h)). \quad (1.17)$$

Then condition $(\sigma 4)$ implies

$$\begin{aligned} |R(t, h, t' + h) - R(t, h, t')| &\leq C 2^{-3(m+1)} (|t - t'| - 2^{-m})^{\gamma-3} \\ &\leq C 2^{-\gamma m} (|i - i'| - 1)^{\gamma-3}. \end{aligned} \quad (1.18)$$

In the same manner one can treat the difference $R(t, h, t' - h) - R(t, h, t')$ and we conclude that

$$|\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m,i'}(\mathbb{G}_1)| \leq C 2^{(1-\gamma)|m|} (1 + |i - i'|)^{\gamma-3}, \quad (1.19)$$

for all $m \geq -1$ and all $i, i' \in J_m$.

Case 2 ($m > m' \geq -1$). First we deal with the case that t is close to t' , $t' - h'$ or $t' + h'$. In this situation we apply again the Cauchy Schwarz inequality and obtain

$$|\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m,i'}(\mathbb{G}_1)| \leq C 2^{(1-\gamma)(m+\overline{m}')/2}.$$

Now we consider those t 's where the distance to the points t' , $t' - h'$ and $t' + h'$ is at least 2^{-m+1} . We distinguish again two different situations.

Case 2.1. Suppose that the interval $[t - h, t + h]$ is contained in the interval $[t' - h', t' + h']$. Using again Taylor's formula we have

$$R(t, h, t' + s) = 2^{-2(m+1)} (D^{(2,0)} \sigma^2(t + \theta h, t' + s) + D^{(2,0)} \sigma^2(t - \theta h, t' + s))$$

for some $\theta \in [0, 1]$ where s can assume the values 0, h' or $-h'$. Thus, condition $(\sigma 4)$ yields

$$|R(t, h, t' + s)| \leq C 2^{-\gamma m} (2^m |t - (t' + s)| - 1)^{\gamma-2}$$

which gives

$$\begin{aligned} |\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m,i'}(\mathbb{G}_1)| &\leq \\ &C 2^{(1-\gamma)\frac{m+\overline{m}'}{2} - \gamma\frac{m-\overline{m}'}{2}} (\min\{|i - (i' + l)2^{m-\overline{m}'}| : l = 0, 1/2, 1\} - 1)^{\gamma-2}. \end{aligned} \quad (1.20)$$

Case 2.2. Now, the interval $[t - h, t + h]$ lies outside $[t' - h', t' + h']$. Let us again consider the difference $R(t, h, t' + h') - R(t, h, t')$ which can be written as telescope sum

$$R(t, h, t' + h') - R(t, h, t') = \sum_{t' \leq l 2^{-m} < t' + h'} R(t, h, (l+1)2^{-m}) - R(t, h, l 2^{-m}).$$

Using (1.17) and (1.18) we deduce

$$|R(t, h, t' + h') - R(t, h, t')| \leq C 2^{-\gamma m} \sum_{t' \leq l 2^{-m} < t' + h'} (|i - l| - 1)^{\gamma-3}.$$

In the same way, one finds an analogous estimate for the difference $R(t, h, t' - h') - R(t, h, t')$ and, concluding, we get for the covariance

$$|\mathbb{E} \Delta^{m,i}(\mathbb{G}_1) \Delta^{m,i'}(\mathbb{G}_1)| \leq C 2^{(1-\gamma)\frac{m+\overline{m}'}{2} - \gamma\frac{m-\overline{m}'}{2}} \sum_{t' - h' \leq l 2^{-m} < t' + h'} (|i - l| - 1)^{\gamma-3}. \quad (1.21)$$

With this preparation let us attack the inequality (1.16). Because of

$$\sum_{|m'|=|m|} \sum_{i' \in J_{m'}} |\mathbb{E} \Delta^{m,i}(\mathbb{G}_d) \Delta^{m',i'}(\mathbb{G}_d)| = \sum_{|m'|=|m|} \prod_{j=1}^d \sum_{i'_j \in J_{m'_j}} |K(\Delta^{m_j,i_j}, \Delta^{m'_j,i'_j})|, \quad (1.22)$$

the problem reduces to the one-dimensional cases considered above. For $m_j = m'_j$, the estimate (1.19) establishes

$$\sum_{i'_j \in J_{m'_j}} |K(\Delta^{m_j,i_j}, \Delta^{m'_j,i'_j})| \leq C 2^{(1-\gamma)\frac{\overline{m}+\overline{m}'}{2}}. \quad (1.23)$$

For $m_j \neq m'_j$, we combine the inequalities (1.20) and (1.21) and obtain

$$\begin{aligned} & \sum_{i'_j \in J_{m'_j}} |K(\Delta^{m_j,i_j}, \Delta^{m'_j,i'_j})| \\ & \leq C 2^{(1-\gamma)\frac{\overline{m}_j+\overline{m}'_j}{2}-\gamma\frac{|\overline{m}_j-\overline{m}'_j|}{2}} \left(\sum_{l=1}^{2^{|\overline{m}_j-\overline{m}'_j|}} l^{\gamma-2} + \sum_{k=0}^{\infty} 2^{|\overline{m}_j-\overline{m}'_j|} \sum_{l=1}^{\infty} (k+l)^{\gamma-3} \right) \\ & \leq C 2^{(1-\gamma)\frac{\overline{m}_j+\overline{m}'_j}{2}} 2^{-(\gamma/2-(\gamma-1)_+)|\overline{m}_j-\overline{m}'_j|}. \end{aligned} \quad (1.24)$$

Observe that $q := \gamma/2 - (\gamma - 1)_+$ is positive for $\gamma \in (0, 2)$. Finally, we substitute (1.23) and (1.24) in (1.22) and we conclude

$$\begin{aligned} \sum_{|m'|=|m|} \sum_{i' \in J_{m'}} |\mathbb{E} \Delta^{m,i}(\mathbb{G}_d) \Delta^{m',i'}(\mathbb{G}_d)| & \leq C 2^{(1-\gamma)|m|} \sum_{|m'|=|m|} 2^{-q \sum_{j=1}^d |\overline{m}_j-\overline{m}'_j|} \\ & \leq C 2^{(1-\gamma)|m|} \left(1 + \sum_{k=1}^{|m|} (2k)^{d-1} 2^{-2qk} \right) \\ & \leq C 2^{(1-\gamma)|m|} \end{aligned}$$

which completes the proof. \square

Corollary 1.8 *Suppose that σ^2 satisfies $(\sigma 1)$ – $(\sigma 4)$ and $K(0, 0) = 0$ then we have for all $m \in \{-1, 0, \dots\}^d$*

$$\sum_{\substack{|m'|=|m| \\ m' \in \{-1, 0, \dots\}^d}} \sum_{\substack{i' \in J_{m'} \\ (m', i') \neq (m, i)}} |\mathbb{E} \Delta^{m,i}(\mathbb{G}_d) \Delta^{m',i'}(\mathbb{G}_d)| \leq C \mathbb{E} \Delta^{m,i}(\mathbb{G}_d)^2 \quad (1.25)$$

for some constant C . If, additionally, we restrict the indices to $m, m' \in \{k, 2k, \dots\}^d$, $i' \in \tilde{J}_{m'} := \prod \{i 2^k : i = 0, \dots, 2^{m'_j-k} - 1\}$ and $i \in \tilde{J}_m$ then we can produce an arbitrary small C in (1.25) by choosing k sufficiently large.

Proof. The estimate (1.25) follows immediately from Lemma 1.7 and Corollary 1.6. The second statement can be checked by a close look to the proof of Lemma 1.7. The modification of the index sets J_m to \tilde{J}_m ensures that (1.23) is as small as desired, while the restriction to $\{k, 2k, \dots\}^d$ makes (1.24) small. \square

1.4 Entropy and Approximation Quantities

Definitions. In this section we want to recall the definitions and properties of some entropy and approximation quantities. We denote by $S : E \rightarrow F$ a bounded operator acting between two Banach spaces. Let B_E and B_F be the unit balls of E and F , respectively. Then the (dyadic) entropy numbers of a compact set $C \subset F$ are defined as

$$e_k(C) := \inf \left\{ \varepsilon > 0 : \exists x_1, \dots, x_{2^{k-1}} \in C \text{ such that } C \subset \bigcup_{j=1}^{2^{k-1}} (x_j + \varepsilon B_F) \right\}$$

and we write $e_k(S)$ instead of $e_k(S(B_E))$. The quantities

$$d_k(C) := \inf \{ \varepsilon > 0 : \exists \tilde{F} \text{ subspace of } F, \dim \tilde{F} < k \text{ and } C \subset \tilde{F} + \varepsilon B_F \}$$

are called Kolmogorov numbers and we use the same convention as above. The operator S is compact iff both quantities tend to zero as k goes to infinity. Next, we introduce the ℓ -norm of an operator S mapping a Hilbert space H into a Banach space. It is defined as

$$\ell(S) := \sup \left\{ \left(\mathbb{E} \left\| \sum_{j=1}^n g_j S f_j \right\|^2 \right)^{1/2} : n \in \mathbb{N}, f_1, \dots, f_n \in H \text{ orthonormal} \right\},$$

where g_1, g_2, \dots are independent $\mathcal{N}(0, 1)$ -distributed random variables. For operators with finite ℓ -norm, we define approximation numbers with respect to the ℓ -norm by

$$\ell_k(S) := \inf \{ \ell(S - S') : S' : H \rightarrow E, \text{rank } S' < k \}.$$

Elementary Properties. The entropy and approximation quantities which were defined above have several properties in common (E and F are reserved for Banach spaces while H shall always denote a Hilbert space)

Monotonicity: we have $\|S\| = e_1(S) \geq e_2(S) \geq \dots \geq e_n(S) \geq \dots$ and, analogously, $\|S\| = d_1(S) \geq d_2(S) \geq \dots$ and $\ell(S) = \ell_1(S) \geq \ell_2(S) \geq \dots$

Additivity: for $S_1, S_2 : E \rightarrow F$ it holds $e_{k+l-1}(S_1 + S_2) \leq e_k(S_1) + e_l(S_2)$ and $d_{k+l-1}(S_1 + S_2) \leq d_k(S_1) + d_l(S_2)$ and for $S_1, S_2 : H \rightarrow E$ we have analogously $\ell_{k+l-1}(S_1 + S_2) \leq \ell_k(S_1) + \ell_l(S_2)$,

Multiplicativity: let $S_1 : E_1 \rightarrow E_2$ and $S_2 : E_2 \rightarrow E_3$ then one can verify $e_{k+l-1}(S_2 S_1) \leq e_k(S_2) e_l(S_1)$ and $d_{k+l-1}(S_2 S_1) \leq d_k(S_2) d_l(S_1)$ and for operators $S_1 : H_1 \rightarrow H_2$ and $S_2 : H_2 \rightarrow E$ it holds $\ell_{k+l-1}(S_2 S_1) \leq \ell_k(S_2) \ell_l(S_1)$,

Surjectivity: let $Q : \tilde{E} \rightarrow E$ be a metric surjection, i.e. $Q(B_{\tilde{E}}) = B_E$, then $e_k(SQ) = e_k(S)$ and $d_k(SQ) = d_k(S)$ and for a metric surjection $Q : \tilde{H} \rightarrow H$ one has $\ell_k(SQ) = \ell_k(S)$.

For the approximation quantities, we have in addition

Rank Property: $d_k(S) = 0$ iff $\text{rank } S < k$ and $\ell_k(S) = 0$ iff $\text{rank } S < k$.

For further details concerning entropy and Kolmogorov numbers we recommend B. Carl and I. Stephani's book [CS90]. For generalized approximation numbers we refer the reader to [Pie87].

Mutual Relations. We can use the following relations between the quantities defined above. Consider again a compact operator S acting from a Hilbert space to a Banach space. Then a theorem of A. Pajor and N. Tomczak-Jaegermann [PTJ86] states that

$$\sup_{k \geq 1} k^{1/2} d_k(S) \leq C \ell(S) \quad (1.26)$$

(see e.g. Pisier [Pis89], Theorem 5.8) and Theorem 1.3 from Carl et al. [CKP] shows that the inequality

$$\sup_{1 \leq k \leq n} b_k e_k(S) \leq C_\alpha \sup_{1 \leq k \leq n} b_k d_k(S) \quad (1.27)$$

holds for all $n \in \mathbb{N}$ provided the sequence $(b_k)_{k \in \mathbb{N}}$ is increasing and satisfies $b_{2k} \leq \alpha b_k$ for some $\alpha \geq 1$ and all $k \in \mathbb{N}$. Taking $b_k = k^{1/2}$ we obtain inequality (1.26) for entropy numbers.

Real and Complex Case. Next, let us mention that the asymptotic behavior of entropy and Kolmogorov numbers is not effected if we change from a real to a complex function space. Let E and F be complex function spaces. We consider each element $f \in E$ as vector $f = (\text{Re}f, \text{Im}f)$ and we have $\|f\| = \|((\text{Re}f)^2 + (\text{Im}f)^2)^{1/2}\|$. In the same manner we deal with F . Thus, we can interpret E and F as spaces over the field of the real numbers. Then we denote by E_r, E_i, F_r and F_i the subspaces of the real and the imaginary parts. We consider the following scheme

$$\begin{array}{ccccc} & & S_r & & \\ & \nearrow Q_{E_i} & E_r & \xrightarrow{\quad} & F_r & \searrow J_{F_r} \\ E & & & & & & F \\ & \searrow Q_{E_r} & E_i & \xrightarrow{\quad} & F_i & \nearrow J_{F_i} \\ & & S_i & & \end{array}$$

where Q_{E_i} and Q_{E_r} denote the quotient maps onto E_r and E_i (which are metric surjections), J_{F_r} and J_{F_i} are the injections, embedding F_r and F_i into F , and $S_r = S_i$. Then the complex version S of the operator S_r can be written as $S = J_{F_r} S_r Q_{E_i} + J_{F_i} S_i Q_{E_r}$. Obviously it holds $d_k(S_r) \leq d_k(S)$. On the other hand, the additivity of the Kolmogorov numbers implies

$$d_{2k-1}(S) \leq d_k(J_{F_r} S_r Q_{E_i}) + d_k(J_{F_i} S_i Q_{E_r}) \leq \|J_{F_r}\| d_k(S_r) + \|J_{F_i}\| d_k(S_i) = 2d_k(S_r).$$

The same arguments work also for entropy numbers.

The spaces $H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ and $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$. From Corollary 1.4 we know that $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ is contained in the RKHS of \mathbb{B}_d^γ , $\gamma = 2\alpha - 1$ and $\alpha \in (1/2, 3/2)$. Let us consider embeddings of the spaces, $H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$, $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ and $\mathcal{H}_{\mathbb{B}_d^{2\alpha-1}}$ into some function space $F([0, 1]^d)$. In the first two cases α may assume an arbitrary value in $(0, \infty)$ while the last case makes sense only for $\alpha \in (1/2, 3/2)$. The subindex zero of $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ stands again for the condition: $f(x_1, \dots, x_d) = 0$ if at least one of the x_j 's equals zero.

Below we want to show that the entropy and Kolmogorov numbers of these embeddings have the same asymptotic behavior. Obviously, we have

$$d_k(\iota : H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)) \leq \begin{cases} d_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)) \\ d_k(\iota : \mathcal{H}_{\mathbb{B}_d^{2\alpha-1}} \rightarrow F([0, 1]^d)) \quad \alpha \in (1/2, 3/2) \end{cases}$$

since $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \subset H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$, for all $\alpha > 0$, and $H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \subset \mathcal{H}_{\mathbb{B}_d^{2\alpha-1}}$, for α in the interval $(1/2, 3/2)$. On the other hand side, we can decompose $H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ and $\mathcal{H}_{\mathbb{B}_d^{2\alpha-1}}$ into a direct sum of 2^d spaces of the form

$$H_a := H_{a_1} \otimes \dots \otimes H_{a_d}, \quad \text{with } a = (a_1, \dots, a_d) \in \{0, 1\}^d,$$

where H_1 stands for $H_{2,0}^\alpha(\mathbb{T})$ while H_0 denotes a one-dimensional space. In the case of $H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ the space H_0 consists of the constant functions and for $\mathcal{H}_{\mathbb{B}_d^{2\alpha-1}}$ we have $H_0 = \text{span}\{K_1^\gamma\}$. Let $|a| = a_1 + \dots + a_d$ and let $\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)$ be a compact embedding of $H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d)$ into some function space $F([0, 1]^d)$. If we have $\|f_0 \otimes f_1\|_{F([0, 1]^{d'})} \leq C\|f_0\|_{F([0, 1])}\|f_1\|_{F([0, 1]^{d'-1})}$ for $f_0 \in H_0$ then we can observe that

$$\begin{aligned} d_{2^d k}(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)) \\ \leq d_k(\iota : H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)) \\ + C \sum_{\substack{a \in \{0, 1\}^d \\ |a| \leq d-1}} d_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^{|a|}) \rightarrow F([0, 1]^{|a|})) \end{aligned}$$

and since

$$d_k(\iota : H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^{d'-1}) \rightarrow F([0, 1]^{d'-1})) \leq d_k(\iota : H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^{d'}) \rightarrow F([0, 1]^{d'})),$$

for all $d' \leq d$, it follows

$$d_{2^d k}(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)) \leq C d_k(\iota : H_{2,0}^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow F([0, 1]^d)).$$

Again, one can apply the same arguments to entropy numbers.

2 Upper Bounds

From the preceding sections one can guess the further strategy. We consider an isometry $T : H \rightarrow \mathcal{H}_{\mathbb{G}_d}$ which maps some Hilbert space onto the RKHS of our process. Then we study T as operator which maps into some Banach space $E \supset \mathcal{H}_{\mathbb{G}_d}$. We search for a convenient decomposition of T into finite dimensional operators R_n . At this point we make use of the Schauder expansions. Next we estimate the ℓ -norm of the R_n 's. Via inequality (1.26) we obtain estimates for the Kolmogorov numbers of the R_n 's. Next, we derive upper bounds for the Kolmogorov numbers of T and inequality (1.27) produces the corresponding entropy bounds. Finally, Theorem 1.1 gives us small ball estimates for most of the investigated norms.

2.1 Key Estimates

Hölder-type Norms. Let $\eta : (0, \infty) \rightarrow (0, \infty)$ denote a continuous, non-decreasing function. We consider the functional

$$\lambda_\eta(f) := \sup_{x \in [0,1]^d} |f(x)| + \sup_{\substack{x, y \in [0,1]^d \\ x \neq y}} \frac{|f(x) - f(y)|}{\eta(|x - y|_2)}, \quad f \in C([0,1]^d),$$

which is a generalization of the Hölder norms. Let us introduce some further notations. We denote by G_n the set $\{i2^{-n} : i = 0, \dots, 2^n\}^d$ and we use the symbol $\kappa_{n,i}$, with $i \in \{0, \dots, 2^n - 1\}^d$, for the dyadic cube $\prod [i_j 2^{-n}, (i_j + 1)2^{-n}]$ of order n . For $f \in C([0,1]^d)$ let

$$D_n(f) := \{|f(x) - f(y)| : x, y \in G_n, |x - y|_2 = 2^{-n}\}.$$

By $C^{\text{lin}}([0,1]^d)$ we denote the 2^d -dimensional space

$$C^{\text{lin}}([0,1]^d) := \text{span}\{f_1 \otimes \dots \otimes f_d : f_j(s) \equiv 1 \text{ or } f_j(s) = s, s \in [0,1]\},$$

the space of all continuous functions on $[0,1]^d$ which are linear in each axe-parallel direction.

Remark that a function $f \in C^{\text{lin}}([0,1]^d)$ attains its minimum and its maximum in the extreme points of $[0,1]^d$, i.e. in the corners $x \in \{0,1\}^d$. In other words $\sup\{|f(x)| : x \in [0,1]^d\} = \max\{|f(x)| : x \in \{0,1\}^d\}$. This can be proved by induction on the dimension d . The same holds true for the derivatives. This implies e.g. that

$$\sup_{x \in [0,1]^d} |D^{(1,0,\dots,0)} f(x)| = \max_{x \in \{0,1\}^{d-1}} |f(0,x) - f(1,x)|. \quad (2.1)$$

Finally, we define

$$C_n^{\text{lin}}([0,1]^d) := \{f \in C([0,1]^d) : \forall i \in \{0, \dots, 2^n - 1\}^d f|_{\kappa_{n,i}} \in C^{\text{lin}}(\kappa_{n,i})\}.$$

Lemma 2.1 *We assume the existence of a constant $c_\eta > 0$ such that η satisfies the inequality $s\eta(t) \leq c_\eta \eta(st)$, for all $s, t \in (0,1)$ (remark that if η is concave on $(0,1)$ this condition is fulfilled with $c_\eta = 1$). Then we have*

$$\lambda_\eta(f) \leq \frac{C}{\eta(2^{-n})} \max\{|f(x)| : x \in G_n\}$$

for all $f \in C_n^{\text{lin}}([0, 1]^d)$, where C depends only on d and c_η .

Proof. Consider two points $x, y \in [0, 1]^d$. We distinguish two cases. Assume first that $|x - y|_2 \geq 2^{-n}$ then the monotonicity of η implies

$$\frac{|f(x) - f(y)|}{\eta(|x - y|_2)} \leq \frac{|f(x) - f(y)|}{\eta(2^{-n})} \leq \frac{2}{\eta(2^{-n})} \max\{|f(x)| : x \in G_n\}.$$

For $|x - y|_2 < 2^{-n}$ the assumption on η yields the inequality

$$\frac{|x - y|_2}{2^{-n}\eta(|x - y|_2)} \leq \frac{c_\eta}{\eta(2^{-n})}$$

and from the observation (2.1) it follows that

$$\frac{|f(x) - f(y)|}{\eta(|x - y|_2)} \leq \frac{|x - y|_2}{\eta(|x - y|_2)} d \frac{\max D_n(f)}{2^{-n}} \leq \frac{2dc_\eta}{\eta(2^{-n})} \max\{|f(x)| : x \in G_n\}$$

which completes the proof. \square

Proposition 2.2 *Suppose that σ^2 satisfies $(\sigma 1)$ for some $\gamma \in (0, 2)$ and let η be as in Lemma 2.1. Then we have the estimate*

$$\mathbb{E} \lambda_\eta((\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d) \leq \frac{C}{\eta(2^{-n})} n^{d/2} 2^{-\gamma n/2},$$

with some constant C depending only on γ , η and d .

Proof. Obviously, we have $(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d \in C_{n+1}^{\text{lin}}([0, 1]^d)$ and by Lemma 2.1 we obtain

$$\begin{aligned} & \mathbb{E} \lambda_\eta((\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d) \\ & \leq \frac{C}{\eta(2^{-(n+1)})} \mathbb{E} \max\{|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)| : x \in G_n\} \\ & \leq \frac{C}{\eta(2^{-n})} \sqrt{1 + \log \#(G_n)} \max\{(\mathbb{E} |(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|^2)^{1/2} : x \in G_n\} \quad (2.2) \end{aligned}$$

(for the second estimate see e.g. Pisier [Pis89], Lemma 4.14 (4.16)). It remains to estimate the variance of

$$(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x) = \sum_{|m|=n} \sum_{i \in J_m} \Delta^{m,i}(\mathbb{G}_d) u_{m,i}(x).$$

Since, for fixed m , the $u_{m,i}$ are essentially disjointly supported, the sum above reduces to a sum over m . Furthermore, we know that $\|u_{m,i}\|_{C([0,1]^d)} \leq 2^{-|m|/2}$. Combining this with Corollary 1.6 we obtain

$$(\mathbb{E} |(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|^2)^{1/2} \leq C \sqrt{\#\{m : |m| = n\}} 2^{-\gamma n/2}$$

(see e.g. (4.15) in Pisier's book [Pis89]). Substitution in (2.2) proves the asserted inequality. \square

Remark. In particular, this shows that the process \mathbb{G}_d possesses a version which is almost surely Hölder continuous of order α (i.e. $\eta(t) = t^\alpha$), for all $\alpha \in [0, \gamma/2)$. This is also true e.g. for $\eta(t) := t^{\gamma/2} (\log t^{-1})^{d/2+1} (\log \log t^{-1})^\beta$, for $t \in (0, \exp(-1))$ and $\beta > 1$.

Now, let us treat a second family of norms in the same manner.

Orlicz Norms. For $2 \leq p < \infty$ consider the Orlicz function $\psi_p(t) := \exp(t^p) - 1$, $t \geq 0$. The Orlicz space $L_{\psi_p}([0, 1]^d)$ consists of all measurable functions f on $[0, 1]^d$ with finite Orlicz norm

$$\|f\|_{\psi_p} := \inf \left\{ c > 0 : \int_{[0, 1]^d} \psi_p(|f(x)|/c) dx \leq 1 \right\}.$$

Then the next result is in complete analogy with Proposition 2.2.

Proposition 2.3 *We assume that σ^2 satisfies $(\sigma 1)$ for some $\gamma \in (0, 2)$. Let $p \in [2, \infty)$. Then we have*

$$\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{\psi_p} \leq C n^{d/2-1/p} 2^{-\gamma n/2},$$

with some constant C depending only on γ , p and d .

Proof. We start with the case $p = 2$. For every fixed $x \in [0, 1]^d$ we have

$$(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x) \stackrel{d}{=} v(x)g,$$

where $g \sim \mathcal{N}(0, 1)$ and $v(x)^2 := \mathbb{E} |(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|^2$. From the proof of Proposition 2.2 we know that $v(x) \leq v := C n^{(d-1)/2} 2^{-\gamma n/2}$. Therefore, for every $t \geq 1$, we have

$$\begin{aligned} \mathbb{E} \psi_2(|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|/2vt) &= \mathbb{E} \exp(g^2 v(x)^2 / 4v^2 t^2) - 1 \\ &\leq \mathbb{E} \exp(g^2 / 4t^2) - 1 \\ &= \left(1 - \frac{1}{2t^2}\right)^{-1/2} - 1 \leq \frac{1}{t^2}. \end{aligned}$$

Integrating x over $[0, 1]^d$ and using Fubini's theorem gives

$$\mathbb{E} \int_{[0, 1]^d} \psi_2(|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|/2vt) dx \leq \frac{1}{t^2}.$$

Whenever $\|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(\cdot, \omega)\|_{\psi_2} > 2vt$, one has by definition of the Orlicz norm

$$\int_{[0, 1]^d} \psi_2(|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x, \omega)|/2vt) dx > 1,$$

hence Čebyshev's inequality yields

$$\mathbb{P}(\omega : \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(\cdot, \omega)\|_{\psi_2} > 2vt) \leq \frac{1}{t^2}.$$

This implies

$$\begin{aligned} \mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{\psi_2} &= \int_0^\infty \mathbb{P}(\omega : \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(\cdot, \omega)\|_{\psi_2} > t) dt \\ &\leq 2v \left(1 + \int_1^\infty \frac{dt}{t^2}\right) = 4v, \end{aligned}$$

i.e. we have shown the assertion for $p = 2$. The case $2 < p < \infty$ can be proved by interpolation. Recall the well known estimate

$$\|f\|_{\psi_p} \leq \|f\|_{C([0,1]^d)}^{1-\theta} \|f\|_{\psi_2}^{\theta}$$

for all $f \in C([0,1]^d)$, where

$$\frac{1}{p} = \frac{1-\theta}{\infty} + \frac{\theta}{2} = \frac{\theta}{2}.$$

Therefore, using Proposition 2.2 and Hölder's inequality we get

$$\begin{aligned} \mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{\psi_p} &\leq \mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{C([0,1]^d)}^{1-\theta} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{\psi_2}^{\theta} \\ &\leq \left(\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{C([0,1]^d)}\right)^{1-\theta} \left(\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d\|_{\psi_2}\right)^{\theta} \\ &\leq C \left(n^{d/2} 2^{-\gamma n/2}\right)^{1-\theta} \left(n^{(d-1)/2} 2^{-\gamma n/2}\right)^{\theta} \\ &\leq C n^{d/2-1/p} 2^{-\gamma n/2} \end{aligned}$$

as asserted. \square

2.2 Entropy and Small Ball Results

Technical Lemma. Before we come to our main arguments, let us prove the following technical lemma.

Lemma 2.4 *Let $S : H \rightarrow E$ be a compact operator and suppose that S is given by $\sum_{n=1}^{\infty} R_n$ (in operator norm) where the operators $R_n : H \rightarrow E$ satisfy*

(R) $\text{rank } R_n \leq a(n)$ and

(L) $\ell(R_n) \leq b(n)$.

Setting

$$a_N := \sum_{n=1}^N a(n) \quad \text{and} \quad b_N := \sum_{n=N+1}^{\infty} b(n),$$

we obtain $d_{2a_N}(S) \leq C a_N^{-1/2} b_N$ for all $N \in \mathbb{N}$.

Proof. By definition of $\ell_k(S)$ we can find an finite rank operator S' with $\text{rank } S' < k$ such that $\ell(S - S') \leq 2\ell_k(S)$. Then additivity and rank property of the Kolmogorov numbers combined with inequality (1.26) imply

$$\begin{aligned} (2k-1)^{1/2} d_{2k-1}(S) &\leq (2k-1)^{1/2} d_k(S - S') + (2k-1)^{1/2} d_k(S') \\ &\leq C \ell(S - S') \leq C \ell_k(S). \end{aligned}$$

Now, we use that S is given by $\sum_{n=1}^{\infty} R_n$. Applying the inequality above with $k = a_N$, we deduce

$$\begin{aligned} d_{2a_N}(S) &\leq C a_N^{-1/2} \ell_{a_N}(S) \\ &\leq C a_N^{-1/2} \ell_{a_N}\left(S - \sum_{n=1}^N R_n\right) \\ &\leq C a_N^{-1/2} \sum_{n=N+1}^{\infty} \ell(R_n) \leq C a_N^{-1/2} b_N \end{aligned}$$

which gives the asserted inequality. \square

Results. In the sequel, many expressions with logarithm will appear. In order to avoid complicate formulas, let us introduce the abbreviation

$$\mathsf{L}t := \max\{1, \log t, \log t^{-1}\}.$$

Now, we are ready to formulate our main theorem concerning Kolmogorov and entropy numbers.

Theorem 2.5 *Denote by (M, μ) some σ -finite measure space. Let T be an integral operator mapping $L_2(M, \mu)$ into $C([0, 1])$. We denote by $k : M \times [0, 1] \rightarrow \mathbb{R}$ its kernel. Let us assume that for some $\gamma \in (0, 2)$ there exists a constant c_γ such that k satisfies the inequality*

$$\int_M (k(x, s) - k(x, t))^2 \mu(dx) \leq c_\gamma |s - t|^\gamma, \quad (2.3)$$

for all $s, t \in [0, 1]$. Then we consider the d -fold tensor product $T_d := T \otimes \cdots \otimes T$ of the operator T , which maps from $L_2(M^d, \mu^{\otimes d})$ into $C([0, 1]^d)$. For $d \geq 2$ and

- for Hölder norms of order $\beta \in [0, \gamma/2)$, i.e. λ_η with $\eta(t) = t^\beta$, we have

$$d_k(T_d) \leq C k^{-(\gamma+1)/2+\beta} (\mathsf{L}k)^{(d-1)((\gamma+1)/2-\beta)+1/2},$$

- for the functional λ_η , with $\eta(t) = t^{\gamma/2} (\mathsf{L}t)^{d/2} (\mathsf{L}^{\circ\nu} t)^\beta \prod_{j=1}^{\nu-1} \mathsf{L}^{\circ j} t$ and $\beta > 1$, we obtain

$$d_k(T_d) \leq C k^{-1/2} (\mathsf{L}^{\circ\nu} k)^{1-\beta},$$

- and for Orlicz norms $\|\cdot\|_{\psi_p}$, with $p \in [2, \infty)$, one has the estimate

$$d_k(T_d) \leq C k^{-(\gamma+1)/2} (\mathsf{L}k)^{(d-1)(\gamma+1)/2+1/2-1/p}.$$

The entropy numbers $e_k(T_d)$ can be estimated from above by the same bounds.

Remark. Let us suppose for example that $k(x, t)$ is uniformly Hölder continuous of order $\gamma/2$ in the second argument and the measure μ is finite. This would be sufficient to ensure the validity of (2.3).

Proof of Theorem 2.5. Setting

$$K(s, t) := \int_M k(x, s)k(x, t) \mu(dx)$$

one can check that K is positive definite. Let $n \in \mathbb{N}$ and take $s_1, \dots, s_n \in [0, 1]$ and choose arbitrary $a_1, \dots, a_n \in \mathbb{R}$, then we have

$$\begin{aligned} \sum_{k,l=1}^n K(s_k, s_l) a_k a_l &= \int_M \sum_{k,l=1}^n a_k a_l k(x, s_k) k(x, s_l) \mu(dx) \\ &= \int_M \left(\sum_{k=1}^n a_k k(x, s_k) \right)^2 \mu(dx) \geq 0. \end{aligned}$$

Consequently, K is a covariance function. Moreover, by assumption it satisfies the condition $(\sigma 1)$ from Section 1.3 and we denote by \mathbb{G}_d again the process with covariance $K^{\otimes d}$ and we obtain immediately a Karhune–Loève expansion

$$\mathbb{G}_d \stackrel{d}{=} \sum_{j=1}^{\infty} g_j T_d f_j$$

where f_1, f_2, \dots forms a complete orthonormal system in $L_2(M, \mu)$ and g_1, g_2, \dots are independent $\mathcal{N}(0, 1)$ -distributed random variables. Then we define R_n to be the composition $(\mathcal{S}^n - \mathcal{S}^{n-1})T_d$ and it follows that

$$\begin{aligned} \ell(R_n) &= \left(\mathbb{E} \left\| \sum_{j=1}^{\infty} g_j (\mathcal{S}^n - \mathcal{S}^{n-1}) T_d f_j \right\|^2 \right)^{1/2} \\ &= \left(\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1}) \mathbb{G}_d\|^2 \right)^{1/2} \\ &\leq C \mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1}) \mathbb{G}_d\|, \end{aligned}$$

where we made use of the equivalence of Gaussian norms (cf. Pisier [Pis89], Corollary 4.9). Then Proposition 2.2 and 2.3 provide us estimates of these ℓ -norms under Hölder-type and Orlicz norms. Hence, the $b(n)$'s in Lemma 2.4 are

$$b(n) = \begin{cases} C n^{d/2} 2^{(\beta-\gamma/2)n} & \text{for the Hölder norms} \\ C(n \prod_{j=1}^{\nu-2} L^{\circ j} n)^{-1} (L^{\circ(\nu-1)} n)^{-\beta} & \text{in the second case} \\ C n^{d/2-1/p} 2^{-\gamma n/2} & \text{for the Orlicz norms} \end{cases}.$$

Then the quantities b_N defined in Lemma 2.4 can be estimated from above by $C N^{d/2} 2^{(\beta-\gamma/2)N}$, $(L^{\circ(\nu-1)} N)^{1-\beta}$ and $C N^{d/2-1/p} 2^{-\gamma N/2}$, respectively. From (1.10) we have $a(n) = C n^{d-1} 2^n$. Thus, one computes easily $a_N \approx C N^{d-1} 2^N$ and we can apply

Lemma 2.4. Let us treat the first case – the Hölder norms – as example. Here, we see that

$$\begin{aligned} b_N &\leq CN^{d/2} 2^{(\beta-\gamma/2)N} = C(N^{d-1} 2^N)^{\beta-\gamma/2} N^{d/2+(d-1)(\gamma/2-\beta)} \\ &\leq Ca_N^{\beta-\gamma/2} (La_N)^{d/2+(d-1)(\gamma/2-\beta)} \end{aligned}$$

and, consequently,

$$d_{2a_N}(T_d) \leq Ca_N^{-(\gamma+1)/2-\beta} (La_N)^{(d-1)((\gamma+1)/2-\beta)+1/2}$$

for all $N \in \mathbb{N}$. Now, consider an arbitrary $k \in \mathbb{N}$ and take N such that k satisfies $2a_N \leq k < 2a_{N+1}$. Obviously, it holds $k \leq Ca_N$. Then the monotonicity of the Kolmogorov numbers yields

$$\begin{aligned} d_k(T_d) &\leq d_{2a_N}(T_d) \leq Ca_N^{-(\gamma+1)/2-\beta} (La_N)^{(d-1)((\gamma+1)/2-\beta)+1/2} \\ &\leq Ck^{-(\gamma+1)/2-\beta} (Lk)^{(d-1)((\gamma+1)/2-\beta)+1/2} \end{aligned}$$

as asserted. The other two cases can be proved in the same manner. The corresponding results for entropy numbers follow from inequality (1.27). \square

Using Theorem 1.1, the small ball estimates below are an immediate consequence of the entropy results above.

Corollary 2.6 *Let \mathbb{G}_d be a centered Gaussian field on $[0, 1]^d$, $d \geq 2$, with covariance $K^{\otimes d}$ where K satisfies*

$$K(s, s) - 2K(s, t) + K(t, t) \leq C|s - t|^\gamma$$

for some $\gamma \in (0, 2)$ and all $s, t \in [0, 1]$.

- For Hölder norms of order $\beta \in [0, \gamma/2)$, one obtains

$$-\log \mathbb{P}(\lambda_{t^\beta}(\mathbb{G}_d) < \varepsilon) \leq C\varepsilon^{-2/(\gamma-2\beta)} (L\varepsilon)^{d-1+d/(\gamma-2\beta)}$$

- and for Orlicz norms $\|\cdot\|_{\psi_p}$, $p \in [2, \infty)$, one has

$$-\log \mathbb{P}(\|\mathbb{G}_d\|_{\psi_p} < \varepsilon) \leq C\varepsilon^{-2/\gamma} (L\varepsilon)^{d-1+d/\gamma-2/(p\gamma)}$$

In particular, this holds for the fractional Brownian sheet \mathbb{B}_d^γ .

Remark. By Theorem 1.1 we cannot expect polynomial bounds like in Corollary 2.6 when we consider a process where the decay of the entropy numbers is of order $k^{-1/2}$ with some additional log-terms. In this situation there is still some work to do.

2.3 More Entropy and Approximation Results

In this section we want to collect all the approximation and entropy results for Bessel potential spaces which follow from Theorem 2.5 and our discussion about the RKHS of the d -dimensional fractional Brownian sheet (see Section 1.1).

More Function Spaces. Let us introduce some more function spaces. First of all, there are the classical spaces of the β -times continuously differentiable, periodic functions, $\beta \in \mathbb{N}$,

$$C^\beta(\mathbb{T}^d) := \{f : \forall m \in \{0, 1, \dots\}^d \text{ with } |m| \leq \beta \ D^m f \in C(\mathbb{T}^d)\}$$

equipped with the norm

$$\|f\|_{C^\beta(\mathbb{T}^d)} := \sum_{|m| \leq \beta} \|D^m f\|_{C(\mathbb{T}^d)}.$$

Now, let $\beta \in (0, \infty)$. Then we denote by $\lfloor \beta \rfloor^-$ the greatest integer which is strictly less than β and we set $\{\beta\}^+ := \beta - \lfloor \beta \rfloor^-$ which, consequently takes values in the interval $(0, 1]$. Assume that $\beta \in (0, \infty) \setminus \mathbb{N}$, then the periodic Hölder spaces of order β are defined as

$$C^\beta(\mathbb{T}^d) := \left\{ f : f \in C^{\lfloor \beta \rfloor^-}(\mathbb{T}^d) \text{ and } \sum_{|m| = \lfloor \beta \rfloor^-} \sup_{\substack{x, y \in \mathbb{T}^d \\ x \neq y}} \frac{|D^m f(x) - D^m f(y)|}{|x - y|_2^{\{\beta\}^+}} < \infty \right\}$$

and

$$\|f\|_{C^\beta(\mathbb{T}^d)} := \|f\|_{C^{\lfloor \beta \rfloor^-}(\mathbb{T}^d)} + \sum_{|m| = \lfloor \beta \rfloor^-} \sup_{\substack{x, y \in \mathbb{T}^d \\ x \neq y}} \frac{|D^m f(x) - D^m f(y)|}{|x - y|_2^{\{\beta\}^+}}.$$

For $\beta \in (0, \infty) \setminus \mathbb{N}$, the periodic Hölder spaces coincide with the periodic Zygmund spaces of order β which are defined as

$$\mathcal{C}^\beta(\mathbb{T}^d) := \left\{ f : f \in C^{\lfloor \beta \rfloor^-}(\mathbb{T}^d) \text{ and } \sum_{|m| = \lfloor \beta \rfloor^-} \sup_{\substack{h \in \mathbb{T}^d \\ h \neq 0}} \frac{\|D^m f(\cdot + 2h) - 2D^m f(\cdot + h) + D^m f\|_{C(\mathbb{T}^d)}}{|h|_2^{\{\beta\}^+}} < \infty \right\}$$

equipped with the norm

$$\|f\|_{\mathcal{C}^\beta(\mathbb{T}^d)} := \|f\|_{C^{\lfloor \beta \rfloor^-}(\mathbb{T}^d)} + \sum_{|m| = \lfloor \beta \rfloor^-} \sup_{\substack{h \in \mathbb{T}^d \\ h \neq 0}} \frac{\|D^m f(\cdot + 2h) - 2D^m f(\cdot + h) + D^m f\|_{C(\mathbb{T}^d)}}{|h|_2^{\{\beta\}^+}}.$$

For $\beta = 1, 2, \dots$ it is obvious by Taylor's formula that the spaces $C^\beta(\mathbb{T}^d)$ embed continuously into the spaces $\mathcal{C}^\beta(\mathbb{T}^d)$, but they are not equivalent. The Zygmund spaces are very convenient for our purposes since they admit equivalent norms which are determined by a sequence of subblocks of the Fourier expansion. Namely, the Zygmund spaces $\mathcal{C}^\beta(\mathbb{T}^d)$ are equivalent to the spaces $B_{\infty, \infty}^\beta(\mathbb{T}^d)$ which are defined as follows. Let $\varphi \in S(\mathbb{R}^d)$ be a non negative function such that $\text{supp } \varphi \subset \{|x|_2 \leq 2\}$ and $\varphi(x) = 1$ for $|x|_2 \leq 1$. We set $\varphi_0 = \varphi$ and we define successively

$$\varphi_n(x) := \varphi(2^{-n}x) - \sum_{j=1}^{n-1} \varphi_j(x) = \varphi(2^{-n}x) - \varphi(2^{-(n-1)}x).$$

We observe that $\text{supp } \varphi_n \subset \{ \lfloor 2^{n-1} \rfloor \leq |x|_2 \leq 2^{n+1} \}$ and that $\sum_{n=0}^{\infty} \varphi_n = \mathbf{1}_{\mathbb{R}^d}$. Then we define for $\beta \geq 0$ the norm

$$\|f\|_{B_{\infty,\infty}^{\beta}(\mathbb{T}^d)}^{\varphi} := \sup_{n=0,1,\dots} 2^{\beta n} \sup_{x \in \mathbb{T}^d} \left| \sum_{k \in \mathbb{Z}^d} \varphi_n(k) \hat{f}(k) \exp(i2\pi kx) \right|$$

and, let $B_{\infty,\infty}^{\beta}(\mathbb{T}^d)$ be the space $\{f \in D'(\mathbb{T}^d) : \|f\|_{B_{\infty,\infty}^{\beta}(\mathbb{T}^d)}^{\varphi} < \infty\}$. One can prove that the norms $\|\cdot\|_{B_{\infty,\infty}^{\beta}(\mathbb{T}^d)}^{\varphi}$ and $\|\cdot\|_{B_{\infty,\infty}^{\beta}(\mathbb{T}^d)}^{\varphi'}$ are equivalent provided φ' satisfies the requirements above. That's why we will omit the superindex φ in the sequel. All these definitions, equivalences and further details can be found in [ST87].

Preparations. Let us consider the operators $L^{\beta} : D'(\mathbb{T}^d) \rightarrow D'(\mathbb{T}^d)$

$$L^{\beta} f := \sum_{k \in \mathbb{Z}^d} \prod_{j=1}^d (1 + |k_j|^2)^{\beta/2} \hat{f}(k) \exp(i2\pi k \cdot)$$

for $\beta \in \mathbb{R}$.

Lemma 2.7 *For $\beta > 0$ the operator $L^{-\beta} : B_{\infty,\infty}^0(\mathbb{T}^d) \rightarrow B_{\infty,\infty}^{\beta}(\mathbb{T}^d)$ is bounded.*

For the proof of this lemma we need the following sublemmata.

Lemma 2.8 *Let $\beta > 0$ and $0 < R < \infty$. We define a symmetric function $P_R^{\beta} : \mathbb{R} \rightarrow \mathbb{R}$ as follows*

$$P_R^{\beta}(t) := \begin{cases} (1 + R^2)^{-\beta/2} \left(1 + \frac{\beta R(R-|t|)}{1+R^2}\right) & |t| \leq R \\ (1 + t^2)^{-\beta/2} & \text{otherwise} \end{cases}.$$

Then we have for the trigonometric polynomial

$$p_R^{\beta}(t) := \sum_{k \in \mathbb{Z}} P_R^{\beta}(k) \exp(i2\pi kt)$$

the estimate $\|p_R^{\beta}\|_{L_1(\mathbb{T})} \leq (1 + \beta)(1 + R^2)^{-\beta/2}$.

Proof. Poisson's summation formula states for $f \in L_1(\mathbb{R})$ that

$$\sum_{k \in \mathbb{Z}} f(t + k) = \sum_{k \in \mathbb{Z}} \exp(i2\pi kt) \int_{\mathbb{R}} f(s) \exp(-i2\pi ks) ds$$

holds for almost all $t \in [0, 1]$. Let us assume that

$$P_R^{\beta}(t) = \int_{\mathbb{R}} f(s) \exp(-i2\pi ts) ds.$$

Since P_R^{β} is symmetric, non negative, convex on the half axes, tending to zero for $t \rightarrow \infty$, Polya's Theorem (see [Fel64], page 509) implies that P_R^{β} is positive definite

and, consequently, by Bochner's Theorem f is the density function of a non negative measure. Thus, we can conclude

$$\begin{aligned} \|p_R^\beta\|_{L_1(\mathbb{T})} &= \int_0^1 \left| \sum_{k \in \mathbb{Z}} f(t+k) \right| dt \\ &= \int_0^1 \sum_{k \in \mathbb{Z}} f(t+k) dt \\ &= \int_{\mathbb{R}} f(t) dt = P_R^\beta(0) \leq (1+\beta)(1+R^2)^{-\beta/2} \end{aligned}$$

which proves our assertion. \square

Lemma 2.9 *Now we consider the difference $\tilde{P}_R^\beta(t) := (1+t^2)^{-\beta/2} - P_R^\beta(t)$. Then we have for the trigonometric polynomial*

$$\tilde{p}_R^\beta(t) := \sum_{k \in \mathbb{Z}} \tilde{P}_R^\beta(k) \exp(i2\pi t)$$

the estimate $\|\tilde{p}_R^\beta\|_{L_1(\mathbb{T})} \leq 1$.

Proof. One can apply the same arguments as in the proof of Lemma 2.8 and obtains $\|\tilde{p}_R^\beta\|_{L_1(\mathbb{T})} \leq \tilde{P}_R^\beta(0) \leq 1$. \square

Proof of Lemma 2.7. Let us consider a fixed $n \in \{0, 1, \dots\}$. We have already remarked that the support of φ_n is contained in the set $\{[2^{n-1}] \leq |x|_2 \leq 2^{n+1}\}$ which is again a subset of $[-2^{n+1}, 2^{n+1}]^d \setminus (-d^{-1/2}[2^{n-1}], d^{-1/2}[2^{n-1}])^d$. Furthermore, we observe that we have the decomposition

$$(1 + |\cdot|^2)^{-\beta/2} = P_{d^{-1/2}[2^{n-1}]}^\beta + \tilde{P}_{d^{-1/2}[2^{n-1}]}^\beta =: P^\beta.$$

Thus, the tensor product $((1 + |\cdot|^2)^{-\beta/2})^{\otimes d}$ can be written as

$$(\tilde{P}_{d^{-1/2}[2^{n-1}]}^\beta)^{\otimes d} + \sum_{j=0}^{d-1} (\tilde{P}_{d^{-1/2}[2^{n-1}]}^\beta)^{\otimes j} \otimes P_{d^{-1/2}[2^{n-1}]}^\beta \otimes (P^\beta)^{\otimes (d-1-j)}.$$

Remark that the first summand is supported by $[-d^{-1/2}[2^{n-1}], d^{-1/2}[2^{n-1}]]^d$. Let us introduce the abbreviations

$$p_n^\beta := \sum_{j=0}^{d-1} (\tilde{p}_{d^{-1/2}[2^{n-1}]}^\beta)^{\otimes j} \otimes p_{d^{-1/2}[2^{n-1}]}^\beta \otimes (p^\beta)^{\otimes (d-1-j)},$$

where $p^\beta := \tilde{p}_{d^{-1/2}[2^{n-1}]}^\beta + p_{d^{-1/2}[2^{n-1}]}^\beta$, and

$$\delta_n(x) := \sum_{k \in \mathbb{Z}^d} \varphi_n(k) \exp(i2\pi kx).$$

Then, we can conclude that

$$\delta_n * (L^{-\beta} f) = p_n^\beta * (\delta_n * f)$$

where $*$ denotes the usual convolution. From Young's inequality and Lemma 2.8 and 2.9 it follows

$$\|\delta_n * (L^{-\beta} f)\|_{C(\mathbb{T}^d)} \leq \|p_n^\beta\|_{L_1(\mathbb{T}^d)} \|\delta_n * f\|_{C(\mathbb{T}^d)} \leq C 2^{-\beta n} \|\delta_n * f\|_{C(\mathbb{T}^d)}$$

and substitution in the definition of the $B_{\infty, \infty}^\beta(\mathbb{T}^d)$ norm yields the boundedness of the operator $L^{-\beta}$. \square

Lemma 2.10 *Let $0 \leq \alpha_1 < \alpha_2$. Then we have*

$$\left. \begin{aligned} d_k(\iota : H_2^{(\alpha_2, \dots, \alpha_2)}(\mathbb{T}^d) &\rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)) \\ e_k(\iota : H_2^{(\alpha_2, \dots, \alpha_2)}(\mathbb{T}^d) &\rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)) \end{aligned} \right\} \approx k^{-(\alpha_2 - \alpha_1)} (\mathbb{L}k)^{(d-1)(\alpha_2 - \alpha_1)}$$

for all $k = 1, 2, \dots$

Proof. Obviously, $L^{\alpha_2 - \alpha_1} : H_2^{(\alpha_2, \dots, \alpha_2)}(\mathbb{T}^d) \rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)$ is an isometry and $(L^{\alpha_2 - \alpha_1})^{-1} = L^{-(\alpha_2 - \alpha_1)}$. Consequently, we have

$$\begin{aligned} d_k(\iota : H_2^{(\alpha_2, \dots, \alpha_2)}(\mathbb{T}^d) &\rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)) \\ &= d_k(L^{-(\alpha_2 - \alpha_1)} : H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d) \rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)) \end{aligned}$$

and, analogously,

$$\begin{aligned} e_k(\iota : H_2^{(\alpha_2, \dots, \alpha_2)}(\mathbb{T}^d) &\rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)) \\ &= e_k(L^{-(\alpha_2 - \alpha_1)} : H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d) \rightarrow H_2^{(\alpha_1, \dots, \alpha_1)}(\mathbb{T}^d)). \end{aligned}$$

Let $\lambda_k(L^{-(\alpha_2 - \alpha_1)})$ denote the k -th eigenvalue of $L^{-(\alpha_2 - \alpha_1)}$. Since we are in the Hilbert space setting we have $d_k(L^{-(\alpha_2 - \alpha_1)}) = \lambda_k(L^{-(\alpha_2 - \alpha_1)})$ (combine i.e. [Pie80], Proposition 11.6.2 and [CS90], Proposition 4.4.1). In addition, from [Car81] we know that for the entropy numbers the inequality $\lambda_k(L^{-(\alpha_2 - \alpha_1)}) \leq 2^{1/2} e_k(L^{-(\alpha_2 - \alpha_1)})$ holds. Next, one can check quickly that $\#\{k \in \mathbb{Z}^d : \prod_{j=1}^d (1 + |k_j|^2)^{-1/2} \geq 2^{-n}\} \approx n^{d-1} 2^n$. Thus, we deduce $\lambda_{n^{d-1} 2^n}(L^{-(\alpha_2 - \alpha_1)}) \approx 2^{-(\alpha_2 - \alpha_1)n}$ and the monotonicity and inequality (1.27) yield the statement of the lemma. \square

Results. We start with embeddings of Bessel potential spaces with dominating mixed smoothness into Zygmund spaces.

Corollary 2.11 *Let $\alpha > 1/2$ then we have*

$$\left. \begin{aligned} d_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow C(\mathbb{T}^d)) \\ e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow C(\mathbb{T}^d)) \end{aligned} \right\} \lesssim k^{-\alpha} (\mathbb{L}k)^{\alpha(d-1)+1/2} \quad (2.4)$$

and, for $0 < \beta < \alpha - 1/2$, it holds

$$\left. \begin{aligned} d_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow \mathcal{C}^\beta(\mathbb{T}^d)) \\ e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow \mathcal{C}^\beta(\mathbb{T}^d)) \end{aligned} \right\} \lesssim k^{-(\alpha - \beta)} (\mathbb{L}k)^{(\alpha - \beta)(d-1)+1/2}.$$

Proof. From Theorem 2.5 and Corollary 1.4, we have (2.4) for $\alpha \in (1/2, 3/2)$. Assume now that $\alpha \geq 3/2$ then the multiplicativity and Lemma 2.10 give

$$\begin{aligned} d_{2k-1}(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow C(\mathbb{T}^d)) \\ &= d_{2k-1}(L^{-(\alpha-1)} : H_2^{(1, \dots, 1)}(\mathbb{T}^d) \rightarrow C(\mathbb{T}^d)) \\ &\leq d_k(L^{-(\alpha-1)} : H_2^{(1, \dots, 1)}(\mathbb{T}^d) \rightarrow H_2^{(1, \dots, 1)}(\mathbb{T}^d)) d_k(\iota : H_2^{(1, \dots, 1)}(\mathbb{T}^d) \rightarrow C(\mathbb{T}^d)) \\ &\leq Ck^{-\alpha}(\mathbb{L}k)^{\alpha(d-1)+1/2} \end{aligned}$$

and, similarly, one deduces the estimate for the entropy numbers.

For the second estimate, we decompose the embedding ι as follows

$$H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \xrightarrow{L^\beta} H_2^{((\alpha-\beta), \dots, (\alpha-\beta))}(\mathbb{T}^d) \xrightarrow{\iota} C(\mathbb{T}^d) \xrightarrow{L^{-\beta}} \mathcal{C}^\beta(\mathbb{T}^d).$$

In this situation L^β is an isometry and, by Lemma 2.7, $L^{-\beta}$ is a bounded operator. Hence, the multiplicativity of Kolmogorov and entropy numbers combined with (2.4) yields the claimed asymptotic bound. \square

In the same manner as we proved estimate (2.4), one can show the following result for our Orlicz norms.

Corollary 2.12 *Let $\alpha > 1/2$ and $p \in [2, \infty)$ then we have*

$$\left. \begin{aligned} d_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow L_{\psi_p}([0, 1]^d)) \\ e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) &\rightarrow L_{\psi_p}([0, 1]^d)) \end{aligned} \right\} \lesssim k^{-\alpha}(\mathbb{L}k)^{\alpha(d-1)+1/2-1/p}.$$

Remark. Let us mention that our result (2.4) improves a result due to V. N. Temlyakov for the Hilbert space case. In [Tem95], he obtained the upper bound

$$e_k(\iota : H_p^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow C(\mathbb{T}^d)) \lesssim k^{-\alpha}(\mathbb{L}k)^{(\alpha+1/2)(d-1)} \quad (2.5)$$

where $p \in (1, \infty)$ (see Theorem 3.3 in [Tem95]). For $d = 2$, there is no difference between (2.4) and (2.5), but for $d \geq 3$ our exponent of the logarithm is smaller.

3 Lower Bounds

3.1 A Lower Estimate for the ℓ -Norm

Let us consider the d -dimensional Brownian sheet under the sup-norm. In this special case we want to show that our key estimate from Proposition (2.2) is sharp.

Proposition 3.1 *For the d -dimensional Brownian sheet, $d \geq 2$, we have*

$$\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1\|_{C([0,1]^d)} \gtrsim n^{d/2} 2^{-n/2}.$$

Proof. We recall that

$$(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1 = \sum_{|m|=n} \sum_{i \in J_m} g_{m,i} u_{m,i}, \quad (3.1)$$

where the $g_{m,i}$'s are independent $\mathcal{N}(0, 1)$ -distributed random variables.

In a first step, we go over to a subprocesses of (3.1). Let $n \in \mathbb{N}$ be sufficiently large, i.e. $n > 4d$, and choose a real number β such that $\beta n \in \mathbb{N}$ and $(4d)^{-1} \leq \beta < (2d)^{-1}$. Then we introduce the index set

$$\begin{aligned} M_n &:= \{m \in \{0, 1, \dots\}^d : |m| = n, m_j \geq \beta n, j = 1, \dots, d\} \\ &= \{(\beta n, \dots, \beta n) + m : |m| = (1 - d\beta)n\}, \end{aligned}$$

which consists of $\binom{(1-d\beta)n+d-1}{d-1}$ elements. The well-known Anderson inequality [And55] states that for any centered Gaussian random variable X with values in a Banach space E and any centrally symmetric, convex set $A \subset E$, it holds

$$\mathbb{P}(X \in A) \geq \mathbb{P}(X \in x + A), \quad (3.2)$$

for all $x \in E$. From this we deduce

$$\begin{aligned} \mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1\|_{C([0,1]^d)} &= \int_0^\infty \mathbb{P}\left(\left\|\sum_{|m|=n} \sum_{i \in J_m} g_{m,i} u_{m,i}\right\|_{C([0,1]^d)} > t\right) dt \\ &\stackrel{(3.2)}{\geq} \int_0^\infty \mathbb{P}\left(\left\|\sum_{m \in M_n} \sum_{i \in J_m} g_{m,i} u_{m,i}\right\|_{C([0,1]^d)} > t\right) dt \\ &= \mathbb{E} \left\|\sum_{m \in M_n} \sum_{i \in J_m} g_{m,i} u_{m,i}\right\|_{C([0,1]^d)}. \end{aligned}$$

Secondly, we observe that all partitions P_m with $m \in M_n$ are finer than $P_{(\beta n, \dots, \beta n)}$, i.e. each element of P_m is a subset of some element of $P_{(\beta n, \dots, \beta n)}$. Since these partitions correspond to the supports of the $u_{m,i}$'s, this yields that the random variables

$$\sum_{m \in M_n} \sum_{i \in J_m} g_{m,i} \mathbf{1}_b u_{m,i},$$

with $b \in P_{(\beta n, \dots, \beta n)}$, are independent.

Next we calculate the average of the variance on $b \in P_{(\beta_n, \dots, \beta_n)}$

$$\begin{aligned}
V_n &:= 2^{d\beta_n} \int_b \mathbb{E} \left(\sum_{m \in M_n} \sum_{i \in J_m} g_{m,i} u_{m,i}(x) \right)^2 dx \\
&= 2^{d\beta_n} \int_b \sum_{m \in M_n} \sum_{i \in J_m} u_{m,i}(x)^2 dx \\
&= 2^{d\beta_n} \#M_n \# \{i \in J_m : \text{supp } u_{m,i} \subset \text{cl}(b)\} 2^{-2n} \int_{[0,1]^d} u_{0,0}(x)^2 dx \\
&\geq C n^{d-1} 2^{-n}.
\end{aligned}$$

Then we choose from each brick $b \in P_{(\beta_n, \dots, \beta_n)}$ a point $x_b \in b$ such that the variance of the random variable

$$g_b := \sum_{m \in M_n} \sum_{i \in J_m} g_{m,i} u_{m,i}(x_b)$$

is at least V_n . From the discussion above we know that the random variables g_b , $b \in P_{(\beta_n, \dots, \beta_n)}$, are independent. Denoting by g_1, g_2, \dots a sequence of independent $\mathcal{N}(0, 1)$ -distributed random variables, we conclude that

$$\begin{aligned}
\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1\|_{C([0,1]^d)} &\geq \mathbb{E} \sup_{x \in [0,1]^d} \left| \sum_{m \in M_n} \sum_{i \in J_m} g_{m,i} u_{m,i}(x) \right| \\
&\geq \mathbb{E} \sup_{b \in P_{(\beta_n, \dots, \beta_n)}} |g_b| \\
&\geq V_n^{1/2} \mathbb{E} \sup_{j=1, \dots, 2^{d\beta_n}} |g_j| \\
&\geq C V_n^{1/2} (\log 2^{d\beta_n})^{1/2} = C n^{d/2} 2^{-n/2},
\end{aligned}$$

where the last estimate follows from the known inequality

$$C(\log N)^{1/2} \leq \mathbb{E} \sup_{j=1, \dots, N} |g_j|$$

(see e.g. [Pis89], Lemma 4.14 (4.18)). This completes the proof. \square

The next lemma gives us a relation between Orlicz norms and the sup-norm.

Lemma 3.2 *Let $f \in C([0, 1]^d)$ be such that the inequality*

$$|f(x) - f(y)| \leq C_1 2^n \|f\|_{C([0,1]^d)} |x - y|_2 \quad (3.3)$$

holds for all $x, y \in [0, 1]^d$. Then we have for the Orlicz norms $\|\cdot\|_{\psi_p}$, $p \in [2, \infty)$,

$$\|f\|_{\psi_p} \geq C_2 n^{-1/p} \|f\|_{C([0,1]^d)},$$

where C_2 depends only on C_1 , p and d .

Proof. Let $x_0 \in [0, 1]^d$ be the point where $|f|$ attains its maximum. Then we can estimate

$$\begin{aligned} & \lambda^d \left\{ x \in [0, 1]^d : |f(x)| \geq \frac{1}{2} \|f\|_{C([0, 1]^d)} \right\} \\ & \geq \lambda^d \left\{ x \in [0, 1]^d : |f(x_0) - f(x)| \leq \frac{1}{2} \|f\|_{C([0, 1]^d)} \right\} \\ & \stackrel{(3.3)}{\geq} \lambda^d \left\{ x \in [0, 1]^d : |x_0 - x| \leq \frac{1}{2C_1 2^n} \right\} \geq C 2^{-dn}. \end{aligned}$$

By definition of the Orlicz norm and Čebyshev's inequality one has

$$2 \geq \int_{[0, 1]^d} \exp\left(\frac{|f(x)|^p}{2^p \|f\|_{\psi_p}^p}\right) dx \geq C 2^{-dn} \exp\left(\frac{\|f\|_{C([0, 1]^d)}^p}{2^p \|f\|_{\psi_p}^p}\right)$$

and taking logarithm and the p -th root gives the asserted inequality. \square

Corollary 3.3 *For the d -dimensional Brownian sheet, $d \geq 2$, under the Orlicz norms $\|\cdot\|_{\psi_p}$, $p \in [2, \infty)$, we have the estimate*

$$\mathbb{E} \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1\|_{\psi_p} \gtrsim n^{d/2-1/p} 2^{-n/2}.$$

Proof. Recall that $(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1 \in C_{n+1}^{\text{lin}}([0, 1]^d)$. Setting $\eta(t) = t$, Lemma 2.1 gives

$$\begin{aligned} & |((\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1)(x) - ((\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1)(y)| \\ & \leq C 2^n \|(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{B}_d^1\|_{C([0, 1]^d)} |x - y|_2. \end{aligned}$$

Combining Proposition 3.1 and Lemma 3.2 we obtain the asserted bound. \square

Hence our estimates from Proposition 2.3 are sharp in the special case where \mathbb{G}_d is the d -dimensional Brownian sheet \mathbb{B}_d^1 .

3.2 Lower Bounds in the Case $d = 2$

In this section, we want to present two techniques which were applied in order to obtain lower bounds for the small ball problem and the entropy numbers, respectively. There is first the method of M. Talagrand which he used in [Tal94] for the two-dimensional Brownian sheet. This method exploits the support properties of modified Schauder expansions. Below, we give a new proof of the main inequality from [Tal94].

The Brownian sheet corresponds to the Bessel potential space $H_{2,0}^{(1,1)}(\mathbb{T}^2)$. One can renorm this space in such a way that the Schauder system forms an orthonormal base. If we turn to spaces $H_{2,0}^{(\alpha,\alpha)}(\mathbb{T}^2)$ with $\alpha \neq 1$ then we loose the chance to have orthogonality for the Schauder system. Here, it is more convenient to deal with the trigonometric system. In a second paragraph below, we present results for trigonometric systems which are due to V. N. Temlyakov.

Talagrand's Method. The problem of finding good lower bounds for entropy numbers of a compact set is always to find large finite dimensional subspaces where the considered set contains a relatively large ball.

“Good Subspaces”. Let us modify the Schauder expansion of the Brownian sheet. For $m \geq 1$, we introduce

$$\tilde{u}_{m-1,i}^0 := \frac{1}{\sqrt{2}}(u_{m,2i} - u_{m,2i+1}) \quad \text{and} \quad \tilde{u}_{m-1,i}^1 := \frac{1}{\sqrt{2}}(u_{m,2i} + u_{m,2i+1})$$

where $i \in J_{m-1}$. Then we extend this to higher dimensions by setting

$$\tilde{u}_{m,i}^a := \tilde{u}_{m_1,i_1}^{a_1} \otimes \cdots \otimes \tilde{u}_{m_d,i_d}^{a_d}$$

for $a \in \{0,1\}^d$, $m \in \{0,1,\dots\}^d$ and $i \in J_m$. Observe that for fixed m and i , the functions $u_{m+(1,\dots,1),2i+a}$, $a \in \{0,1\}^d$, and $\tilde{u}_{m,i}^a$, $a \in \{0,1\}^d$, span the same subspace and the coefficients of the two expansions

$$\sum_{a \in \{0,1\}^d} c_a u_{m+(1,\dots,1),2i+a} = \sum_{a \in \{0,1\}^d} c'_a \tilde{u}_{m,i}^a$$

are linked via a unitary matrix. Thus we can conclude that

$$\begin{aligned} \mathbb{B}_d^1 &\stackrel{d}{=} \sum_{m \in \{-1,0,\dots\}^d} \sum_{i \in J_m} g_{m,i} u_{m,i} \\ &\stackrel{d}{=} \sum_{m \in \{-1,0,\dots\}^d \setminus \{1,2,\dots\}^d} \sum_{i \in J_m} g_{m,i} u_{m,i} + \sum_{a \in \{0,1\}^d} \sum_{m \in \{0,1,\dots\}^d} \sum_{i \in J_m} g_{m,i}^a \tilde{u}_{m,i}^a \end{aligned}$$

where the $g_{m,i}^a$'s are again independent $\mathcal{N}(0,1)$ -distributed random variables. This follows from a special property of the normal distribution. Namely, let (g_1, \dots, g_n) be a vector of independent $\mathcal{N}(0,1)$ -distributed random variables and let U be a unitary matrix then the vector $(g'_1, \dots, g'_n)^t = (g_1, \dots, g_n)U$ has again independent $\mathcal{N}(0,1)$ -distributed components.

Next, Anderson's inequality (3.2) shows that we can go over to the investigation of a subprocess, namely, we have

$$\mathbb{P}(\|\mathbb{B}_d^1\| < \varepsilon) \leq \mathbb{P}\left(\left\|\sum_{m \in M} \sum_{i \in J_m} g_{m,i} \tilde{u}_{m,i}^{(0,\dots,0)}\right\| < \varepsilon\right)$$

for all $M \subset \{0,1,\dots\}^d$. We chose the functions $\tilde{u}_{m,i}^{(0,\dots,0)}$ since they have the advantage that they assume both positive and negative values.

“Large Balls”. For fixed m , let us assemble all functions $g_{m,i} \tilde{u}_{m,i}^{(0,\dots,0)}$

$$\tilde{B}_m := \sum_{i \in J_m} g_{m,i} \tilde{u}_{m,i}^{(0,\dots,0)}$$

and define $\xi_m(x) := \text{sign}(\tilde{B}_m(x))$. Obviously, we have for almost all ω (none of the $g_{m,i}$'s equals zero)

$$\lambda^d\{x : \xi_m(x, \omega) = 1\} = \lambda^d\{x : \xi_m(x, \omega) = -1\} = \frac{1}{2}$$

and, thus, we can interpret $\xi_m(\cdot, \omega)$ as Bernoulli variable. If we knew that $\xi_m(\cdot, \omega)$, $m \in M$, were independent then we could compute easily the measure of the set of points x where the signs of $\tilde{B}_m(x, \omega)$ are the same for all $m \in M$ and, hence, $\sum_M \tilde{B}_m(x, \omega)$ assumes “large” values. In addition, we wish that all $g_{m,i} \tilde{u}_{m,i}^{(0,\dots,0)}$ are involved in this production of “large” values. Therefore we need a certain “local” independence. The following definition makes these ideas precise.

Definition 3.1 We say that $M \subset \{0, 1, \dots\}^d$ possesses the **local independence property** if for each $m \in M$ the following condition is satisfied. Let $m^* := m + (3, \dots, 3)$ and let b be an arbitrary brick from the partition P_{m^*} defined in (1.7). Then we consider the probability measure $(\lambda^d(b))^{-1} \lambda^d(\cdot \cap b)$ and require that the Bernoulli variables $\xi_{m'}(\cdot, \omega) \mathbf{1}_b$, $m' \in M \setminus \{m\}$, are independent for almost all ω (none of the $g_{m,i}$'s equals zero).

In addition it is useful to collect such m of the same order, since in this case the norms $\|\tilde{u}_{m,i}^{(0,\dots,0)}\|_{C([0,1]^d)} = 2^{-(2d+|m|/2)}$ are all the same. Then Talagrand's result can be formulated as follows.

Proposition 3.4 (M. Talagrand) *Let $M \subset \{m \in \{0, 1, \dots\}^d : |m| = n\}$ and suppose that M possesses the local independence property. Then we have*

$$\sup_{x \in [0,1]^d} \left| \sum_{m \in M} \tilde{B}_m(x, \omega) \right| \geq 2^{-4d} 2^{-3n/2} \sum_{m \in M} \sum_{i \in J_m} |g_{m,i}(\omega)| \quad (3.4)$$

Proof. First, we assume that none of the $g_{m,i}(\omega)$'s equals zero. It is obvious that the local independence property implies the independence of ξ_m , $m \in M$ (again with respect to the probability measure $\lambda^d(\cdot \cap [0, 1]^d)$). Thus, the Lebesgue measure of the set

$$A(\omega) := \{x : \forall m \in M \ \xi_m(x, \omega) = 1\}$$

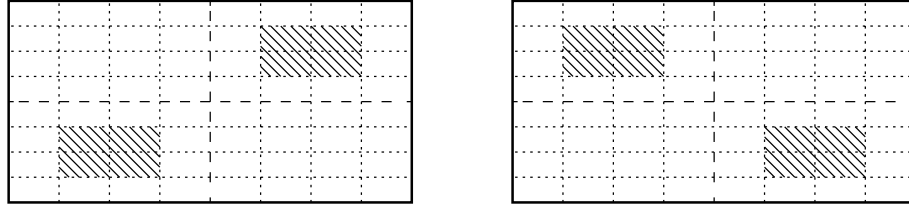
equals $2^{-\#(M)}$. Next we study the sets

$$\begin{aligned} A_{m,i}(\omega) &:= \{x : g_{m,i}(\omega) \tilde{u}_{m,i}^{(0,\dots,0)}(x) \geq 2^{-d} \|g_{m,i}(\omega) \tilde{u}_{m,i}^{(0,\dots,0)}\|_{C([0,1]^d)} \\ &\quad \text{and } \forall m' \in M \setminus \{m\} \ \xi_{m'}(x, \omega) = 1\}. \end{aligned}$$

Recall that $m^* = m + (3, \dots, 3)$. Then it follows that $\text{supp } \tilde{u}_{m,i}^{(0,\dots,0)}$ contains 2^{3d} bricks from P_{m^*} . One can check that the condition

$$g_{m,i}(\omega) \tilde{u}_{m,i}^{(0,\dots,0)}(x) \geq 2^{-d} \|g_{m,i}(\omega) \tilde{u}_{m,i}^{(0,\dots,0)}\|_{C([0,1]^d)} \quad (3.5)$$

is satisfied on 2^{2d-1} of these bricks in all points. The pictures below illustrates the two possible situations for $d = 2$.



For $g_{m,i}(\omega) > 0$, the shaded bricks in the left picture satisfy (3.5) and in the right picture we shaded the bricks where (3.5) holds if $g_{m,i}(\omega) < 0$. Combining this with the local independence property we obtain

$$\lambda^d(A_{m,i}(\omega)) \geq 2^{2d-1} \lambda^d(b_{m^*}) 2^{-(\#(M)-1)} = 2^{-d} 2^{-n} 2^{-\#(M)}.$$

Now, it remains to assemble all these facts. We can conclude

$$\begin{aligned} \sup_{x \in [0,1]^d} \left| \sum_{m \in M} \tilde{B}_m(x, \omega) \right| &\geq \frac{1}{\lambda^d(A(\omega))} \int_{A(\omega)} \sum_{m \in M} \tilde{B}_m(x, \omega) dx \\ &= 2^{\#(M)} \sum_{m \in M} \sum_{i \in J_m} \int_{A(\omega)} g_{m,i}(\omega) \tilde{u}_{m,i}^{(0,\dots,0)}(x) dx \\ &\geq 2^{\#(M)} \sum_{m \in M} \sum_{i \in J_m} 2^{-d} \|g_{m,i}(\omega) \tilde{u}_{m,i}^{(0,\dots,0)}\|_{C([0,1]^d)} \lambda^d(A_{m,i}(\omega)) \\ &\geq 2^{-4d} 2^{-3n/2} \sum_{m \in M} \sum_{i \in J_m} |g_{m,i}(\omega)| \end{aligned}$$

as claimed in (3.4). If some of the $g_{m,i}$'s are zero then the continuity of the functionals on both sides of the inequality yields (3.4). \square

Lemma 3.5 *The sets $M_{3n} := \{m \in \{3k : k = 0, 1, \dots\}^2 : |m| = 3n\}$ have the local independence property.*

Proof. Recall that Bernoulli variables $\varepsilon_1, \dots, \varepsilon_k$ are independent if and only if for each set $J \in \{1, \dots, k\}$ it holds $\mathbb{E} \prod_{i \in J} \varepsilon_i = 0$.

Let us fix a $m \in M_{3n}$ and choose an arbitrary set $J \subset M_{3n} \setminus \{m\}$. Let us sort the elements from $J := \{m^1, \dots, m^k\}$ in such a way that $m_1^1 < m_1^2 < \dots < m_1^k$. Since $m_1^i + m_2^i = 3n$ it follows that $m_2^1 > m_2^2 > \dots > m_2^k$. Assume that for all $i = 1, \dots, k$ we have $m_1 > m_1^i$ and $m_2 > m_2^i$. Summation implies that $3n > m_1^i + m_2^i$ which gives a contradiction. Hence, we may assume without loss of generality that $m_1 < m_1^k$ and, consequently, $m_1^* \leq m_1^k$, i.e. for fixed x_2 the sets $(l2^{-m_1^k}, (l+1)2^{-m_1^k}) \times \{x_2\}$ are either subsets of $b \in P_{m^*}$ or they do not intersect with b . Now, observe that for each fixed x_2 the functions $\xi_{m^i}((\cdot, x_2), \omega)$, $i = 1, \dots, k-1$, are constant on the dyadic intervals of length $2^{-m_1^k}$ while $\xi_{m^k}((\cdot, x_2), \omega)$ takes the value -1 on one half of the interval and 1 on the other half. Thus, we have for $b \in P_{m^*}$

$$\begin{aligned}
& \mathbb{E} \prod_{i=1}^k \xi_{m^i}(\cdot, \omega) \mathbf{1}_b \\
&= \frac{1}{\lambda^2(b)} \int_{\mathbb{R}^2} \mathbf{1}_b(x) \prod_{i=1}^k \xi_{m^i}(x, \omega) dx \\
&= \frac{1}{\lambda^2(b)} \int_{\mathbb{R}} \sum_{l=0}^{2^{m_1^k}-1} \prod_{i=1}^{k-1} \xi_{m^i} \left(\left(\left(l + \frac{1}{2} \right) 2^{-m_1^k}, x_2 \right), \omega \right) \int_{\mathbb{R}} \mathbf{1}_b(x) \xi_{m^k}((x_1, x_2), \omega) dx_1 dx_2 \\
&= 0
\end{aligned}$$

which proves our assertion. \square

Remark. The sets M_{3n} have cardinality $n + 1$ and this combined with the inequality (3.4) will proof the sharpness of the small ball estimate for the sup-norm in the case $d = 2$ as we shall see in a second.

If one wished to use the same method in order to show that our estimates for higher dimensions are sharp too, then one would need sets $M_n^d \subset \{|m| = n\}$ with cardinality of order $n^{(2d-1)/3}$. Yet, this is impossible. Observe, that the partition $P_{(n,\dots,n)}$ is finer than all partitions P_m , $|m| = n$, and on the elements of $P_{(n,\dots,n)}$ the vector $(\xi_m)_{|m|=1}$ is constant. Hence, $(\xi_m)_{|m|=1}$ cannot assume more than 2^{dn} values, but a vector of $Cn^{(2d-1)/3}$ independent Bernoulli variables assumes $2^{Cn^{(2d-1)/3}}$ different values. That's why one has to seek for new ideas for the dimensions $d \geq 3$.

Results. Then one obtains the following lower bounds

Theorem 3.6 *For the two-dimensional Brownian sheet we have*

$$-\log \mathbb{P}(\|\mathbb{B}_2^1\|_{C([0,1]^2)} < \varepsilon) \geq C\varepsilon^2(L\varepsilon)^3$$

and, for $p \in [2, \infty)$,

$$-\log \mathbb{P}(\|\mathbb{B}_2^1\|_{\psi_p} < \varepsilon) \geq C\varepsilon^2(L\varepsilon)^{3-2/p}.$$

Proof. Let g_1, \dots, g_k be independent $\mathcal{N}(0, 1)$ -distributed random variables. We use that there exists a constant C such that the inequality

$$\mathbb{P}\left(\frac{1}{k} \sum_{l=i}^k |g_l| < C\right) \leq \exp(-Ck) \tag{3.6}$$

holds for all $k = 1, 2, \dots$ (apply Čebychev's inequality with the convex function $\exp(-C\cdot)$ and choose $C = -\log(\mathbb{E} \exp(-|g|))/2$). By Proposition 3.4 we can estimate

$$-\log \mathbb{P}(\|\mathbb{B}_2^1\|_{C([0,1]^2)} < \varepsilon)$$

$$\begin{aligned}
&\geq -\log \mathbb{P} \left(\sup_{x \in [0,1]^2} \left| \sum_{m \in M_{3n}} \tilde{B}_m(x) \right| < \varepsilon \right) \\
&\geq -\log \mathbb{P} \left(\frac{1}{(n+1)2^{3n}} \sum_{m \in M_{3n}} \sum_{i \in J_m} |g_{m,i}| < 2^8(n+1)2^{3n/2}\varepsilon \right) \\
&\stackrel{(3.6)}{\geq} C(n+1)2^{3n}
\end{aligned}$$

for $\varepsilon = C2^{-8}(n+1)^{-1}2^{-3n/2}$ where C is the constant from (3.6). Using the fact that $\sum_{m \in M_{3n}} \tilde{B}_m$ is in $C_{3n+2}^{\text{lin}}([0,1]^2)$ we can apply first Lemma 2.1 and then Lemma 3.2 and obtain

$$\begin{aligned}
&-\log \mathbb{P}(\|\mathbb{B}_2^1\|_{\psi_p} < \varepsilon) \\
&\geq -\log \mathbb{P} \left(\frac{1}{(n+1)2^{3n}} \sum_{m \in M_{3n}} \sum_{i \in J_m} |g_{m,i}| < C'n^{1-1/p}2^{3n/2}\varepsilon \right) \\
&\stackrel{(3.6)}{\geq} C(n+1)2^{3n}
\end{aligned}$$

for $\varepsilon = CC'^{-1}n^{1/p-1}2^{-3n/2}$. Then the usual monotonicity arguments (see the proof of Theorem 2.5) complete the proof. \square

By Theorem 1.1 one can deduce the corresponding lower bounds of entropy and Kolmogorov numbers for the embeddings of the space $H_2^{(1,1)}(\mathbb{T}^2)$ into $C(\mathbb{T}^2)$ and $L_{\psi_p}([0,1]^2)$, respectively.

Temlyakov's Method. First we need some further notations. For $m \in \{0, 1, \dots\}$, denote by $B^1(m)$ the dyadic block of integers

$$\{k \in \mathbb{Z} : \lfloor 2^{m-1} \rfloor \leq |k| < 2^m\}$$

and for $m = (m_1, m_2) \in \{0, 1, \dots\}^2$, we set $B^2(m) := B^1(m_1) \times B^1(m_2)$. Then we define the set $F(2n) \subset \mathbb{Z}^2$ to be

$$F(2n) := \bigcup_{\substack{m \in \{2k : k=0,1,\dots\}^2 \\ m_1+m_2=2n}} B^2(m)$$

which forms a kind of hyperbolic cross in the plane. Let G be a subset of \mathbb{Z}^2 then we denote by $T(G)$ the trigonometric polynomials f of the form

$$f(s, t) = \sum_{k \in G} a_k \exp(i(k_1 s + k_2 t)).$$

The next theorem from Temlyakov [Tem95] gives us lower bounds for embeddings of Bessel potential spaces into $C([0,1]^2)$.

Theorem 3.7 (V. N. Temlyakov) *Let $\alpha > 1/2$, there exist constants C, C_∞ and C_α such that for all $n \in \mathbb{N}$ one can find a set of functions $\{f_1, \dots, f_{K(n)}\} \subset T(F(2n))$ which has the following properties*

- (a) $\log K(n) \geq Cn2^{2n}$,
- (b) $\|f_k - f_l\|_{C([0,1]^2)} \geq C_\infty n$, whenever $k \neq l$,
- (c) and $\|f_k\|_{H_2^{(\alpha,\alpha)}(\mathbb{T}^2)} \leq C_\alpha n^{1/2} 2^{2n\alpha}$, for all k .

Consequently, one obtains for the embedding ι

$$\left. \begin{aligned} e_k(\iota : H_2^{(\alpha,\alpha)}(\mathbb{T}^2) &\rightarrow C([0,1]^2)) \\ d_k(\iota : H_2^{(\alpha,\alpha)}(\mathbb{T}^2) &\rightarrow C([0,1]^2)) \end{aligned} \right\} \gtrsim k^{-\alpha} (\mathbb{L}k)^{\alpha+1/2}$$

for all $k \in \mathbb{N}$ (the estimate for the Kolmogorov numbers follows from (1.27) combined with the upper bounds from Corollary 2.11).

Corollary 3.8 *For the entropy numbers of the embedding of $H_2^{(\alpha,\alpha)}(\mathbb{T}^2)$ into the Orlicz space $L_{\psi_p}([0,1]^2)$, $p \in [2, \infty)$, we have*

$$\left. \begin{aligned} e_k(\iota : H_2^{(\alpha,\alpha)}(\mathbb{T}^2) &\rightarrow L_{\psi_p}([0,1]^2)) \\ d_k(\iota : H_2^{(\alpha,\alpha)}(\mathbb{T}^2) &\rightarrow L_{\psi_p}([0,1]^2)) \end{aligned} \right\} \gtrsim k^{-\alpha} (\mathbb{L}k)^{\alpha+1/2-1/p}$$

for all $k \in \mathbb{N}$.

The proof of this corollary is based on the next lemma.

Lemma 3.9 *The functions $f \in T(\{-2^n, \dots, 2^n\}^2)$ satisfy the condition*

$$|f(x) - f(y)| \leq C2^n \|f\|_{C([0,1]^d)} |x - y|_2$$

for all $x, y \in [0,1]^d$.

Proof. We denote by F_n the Fejér kernel

$$F_n(t) := \sum_{|k| \leq 2^n} \left(1 - \frac{|k|}{2^n}\right) \exp(i2\pi kt)$$

and let $V_n(t) := 2F_{n+1}(t) - F_n$ be the de la Vallée-Poussin kernel. Then we define

$$D_n(t) := i2\pi 2^n \exp(i2\pi 2^n t) F_n(t) - i2\pi 2^n \exp(-i2\pi 2^n t) F_n(t).$$

Observe that the k -th coefficient of the Fourier expansion of D_n equals $i2\pi k$, for all k with $|k| \leq 2^n$. Hence, for any trigonometric polynomial $f \in T(\{-2^n, \dots, 2^n\})$ we have

$$D_n * f(t) = \frac{d}{dt} f(t)$$

where $*$ denotes the usual convolution. Recall, that $\|F_n\|_{L_1([0,1])} = 1$ for all $n \in \mathbb{N}$ (see e.g. Temlyakov [Tem93]) which yields $\|D_n \otimes V_n\|_{L_1([0,1]^2)} \leq 12\pi 2^n$ for all $n \in \mathbb{N}$. By Young's inequality it follows

$$\|D^{(1,0)} f\|_{C([0,1]^2)} = \|(D_n \otimes V_n) * f\|_{C([0,1]^2)} \leq C2^n \|f\|_{C([0,1]^2)}$$

for all $f \in T(\{-2^n, \dots, 2^n\}^2)$. Analogously, one can treat $D^{(0,1)}f$ and we deduce that

$$|f(x) - f(y)| \leq C2^n \|f\|_{C([0,1]^2)} |x - y|_2$$

for all $f \in T(\{-2^n, \dots, 2^n\}^2)$ and all $x, y \in [0, 1]^2$. \square

Proof of Corollary 3.8. Taking the same set of functions as in Theorem 3.7 it follows from Lemma 3.9 and 3.2 that we can replace (b) by

$$(b') \quad \|f_k - f_l\|_{\psi_p} \geq Cn^{1-1/p}, \text{ whenever } k \neq l,$$

and this yields

$$e_{Cn2^{2n}}(\iota : H_2^{(\alpha, \alpha)}(\mathbb{T}^2) \rightarrow L_{\psi_p}([0, 1]^2)) \geq Cn^{1/2-1/p} 2^{-2n\alpha}.$$

Finally, the monotonicity of the entropy numbers implies the desired lower bound and the result for the Kolmogorov numbers follows from inequality (1.27) combined with the upper bounds from Corollary 2.12. By (1.27) with $b_k = k^{\alpha+1}$ there exists a \tilde{k} in $\{1, \dots, k\}$ such that

$$Ck(Lk)^{\alpha+1/2-1/p} \leq \tilde{k}^{\alpha+1} d_{\tilde{k}}(\iota : H_2^{(\alpha, \alpha)}(\mathbb{T}^2) \rightarrow L_{\psi_p}([0, 1]^2)) \leq C'\tilde{k}(L\tilde{k})^{\alpha+1/2-1/p},$$

where the last estimate follows from Corollary 2.12. Consequently, there exists a constant $C \in (0, 1)$ such that we have always $Ck \leq \tilde{k}$. Then the monotonicity of the Kolmogorov numbers completes the proof. \square

Corollary 3.10 *Let us consider the fractional Brownian sheet \mathbb{B}_2^γ , $\gamma \in (0, 2)$, on the unit square. For Orlicz norms $\|\cdot\|_{\psi_p}$, $p \in [2, \infty)$, and sup-norm the small ball estimates from Corollary 2.6 are asymptotically sharp, i.e. the inverse relation “ \gtrsim ” is also true.*

Proof. By Corollary 1.4 we know that $H_{2,0}^{(\alpha, \alpha)}(\mathbb{T}^2)$ is contained in the RKHS of \mathbb{B}_2^γ , where $\alpha = (\gamma + 1)/2$. From the discussion in Section 1.1 we know that for embeddings of $H_{2,0}^{(\alpha, \alpha)}(\mathbb{T}^2)$ we have the same lower bounds as we have in Theorem 3.7 and Corollary 3.8 for the corresponding embeddings of $H_2^{(\alpha, \alpha)}(\mathbb{T}^2)$. Then Theorem 1.1 establishes lower bounds of the same order as the already proved upper ones. \square

It remains open whether one can show similar bounds for the embeddings into Hölder spaces.

3.3 Lower Bounds for $d > 2$.

For $d > 2$ we were not able to prove lower bounds which would show the sharpness of our results from Theorem 2.5. We had already mentioned that

$$e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow C([0, 1]^d)) \leq Ck^{-\alpha}(Lk)^{\alpha(d-1)+1/2}.$$

improves V. N. Temlyakov's result (2.5) in the Hilbert space case. Using Temlyakov's lower bounds for the embedding into $L_1([0, 1]^d)$ (see [Tem90], Theorem 2.4) we obtain

$$\begin{aligned} k^{-\alpha}(\mathsf{L}k)^{\alpha(d-1)} &\lesssim e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow L_1([0, 1]^d)) \\ &\lesssim e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow L_{\psi_p}([0, 1]^d)) \\ &\lesssim e_k(\iota : H_2^{(\alpha, \dots, \alpha)}(\mathbb{T}^d) \rightarrow C(\mathbb{T}^d)), \end{aligned}$$

and we see that the error in the exponent of the log-term can be at most $1/2$ (and, consequently, at most 1 for the small ball estimates of fractional Brownian sheet), independently of the dimension $d > 2$. Moreover, our result for the Orlicz norm $\|\cdot\|_{\psi_2}$ is sharp. Again, we didn't find similar results for Hölder spaces.

Appendix

First, we show the following equivalence of norms.

Proposition A.1 *Let $\varphi \in S(\mathbb{R})$ with $\text{supp } \varphi \subset [0, 1]$ and $\alpha \in (1/2, 3/2)$. Then there exist constants depending on α such that*

$$C_1 \|\varphi\|_{H_2^s(\mathbb{T})} \leq \|\varphi\|_{H_2^s((0,1))} \leq C_2 \|\varphi\|_{H_2^s(\mathbb{T})}. \quad (\text{A.1})$$

For the proof of Proposition A.1 we need the concept of maximal functions. Let f be a locally integrable function on \mathbb{R} . Then the Hardy–Littlewood maximal function is defined by

$$(Mf)(t) := \sup_{r>0} \frac{1}{2r} \int_{t-r}^{t+r} |f(s)| ds$$

for $t \in \mathbb{R}$. The famous Hardy–Littlewood inequality states that for each $p \in (1, \infty]$ there exists a constant C such that

$$\|Mf\|_{L_p(\mathbb{R})} \leq C \|f\|_{L_p(\mathbb{R})} \quad (\text{A.2})$$

for all $f \in L_p(\mathbb{R})$. For a proof and further details we refer the reader to the first chapter in [Ste70].

Next, let us introduce the abbreviation $\rho_\alpha(t) := (1+|t|^2)^{\alpha/2}$. Using the inequality $2(1+a^2)(1+b^2) \geq 1+(a+b)^2$, one can check that

$$\rho_\alpha(s) \leq 2^{|\alpha|/2} (1+|s-t|^2)^{|\alpha|/2} \rho_\alpha(t) \quad (\text{A.3})$$

and, consequently,

$$C \rho_\alpha(s) \leq \rho_\alpha(t) \leq C' \rho_\alpha(s) \quad (\text{A.4})$$

for all $s, t \in \mathbb{R}$ with $|s-t| \leq 2\pi$.

We continue with some preparatory lemmata.

Lemma A.2 *Let $\varphi \in S(\mathbb{R})$ be such that $\text{supp } \mathcal{F}^{-1}\varphi \subset [0, 1]$. Then there exists a constant C depending on α such that*

$$\sup_{s \in \mathbb{R}} \frac{\rho_\alpha(t-s) |D\varphi(t-s)|}{1+|s|} \leq C \sup_{s \in \mathbb{R}} \frac{\rho_\alpha(t-s) |\varphi(t-s)|}{1+|s|}$$

for all $t \in \mathbb{R}$.

Proof. Let $\tilde{\psi} \in S(\mathbb{R})$ be such that $\tilde{\psi}|_{[0,1]} \equiv 1$ and denote by ψ its Fourier transform $\mathcal{F}\tilde{\psi}$. Then it follows that $\mathcal{F}^{-1}\varphi = \mathcal{F}^{-1}\varphi \mathcal{F}^{-1}\psi$ and applying the Fourier transform to both sides yields

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(s) \psi(t-s) ds$$

Thus, using (A.3) we can estimate

$$\begin{aligned} \rho_\alpha(t)|D\varphi(t)| &\leq C \int_{\mathbb{R}} \rho_\alpha(u)|\varphi(u)|(1+|t-u|^2)^{|\alpha|/2}|D\psi(t-u)| du \\ &\leq C \int_{\mathbb{R}} \rho_\alpha(u)|\varphi(u)|(1+|t-u|^2)^{-2} du, \end{aligned} \quad (\text{A.5})$$

since $\psi \in S(\mathbb{R})$.

Next, let us quickly check that the inequality

$$1 + |a| \leq (1 + |b|)(1 + |a - b|) \quad (\text{A.6})$$

holds for all $a, b \in \mathbb{R}$. For $|a| \leq |b|$, this is obvious. Assume now that $|b| = \theta|a|$ with $\theta \in [0, 1)$. Then we have

$$(1 + |b|)(1 + |a - b|) \geq (1 + \theta|a|)(1 + (1 - \theta)|a|) \geq 1 + (\theta + (1 - \theta))|a| = 1 + |a|,$$

which verifies (A.6).

Finally, we substitute in (A.5) t by $t - s$ and divide both sides by $1 + |s|$. Using inequality (A.6) we obtain

$$\begin{aligned} \frac{\rho_\alpha(t-s)|D\varphi(t-s)|}{1 + |s|} &\leq C \int_{\mathbb{R}} \frac{\rho_\alpha(u)|\varphi(u)|}{1 + |t-u|} (1 + |t-s-u|)(1 + |t-s-u|^2)^{-2} du \\ &\leq C \int_{\mathbb{R}} \frac{\rho_\alpha(t-v)|\varphi(t-v)|}{1 + |v|} (1 + |v-s|^2)^{-1} dv \\ &\leq C \sup_{v \in \mathbb{R}} \frac{\rho_\alpha(t-v)|\varphi(t-v)|}{1 + |v|}, \end{aligned}$$

which proves the assertion. \square

Lemma A.3 Consider a function $\varphi \in S(\mathbb{R})$ which satisfies $\text{supp } \mathcal{F}^{-1}\varphi \subset [0, 1]$. Then we have

$$\sup_{s \in \mathbb{R}} \frac{\rho_\alpha(t-s)|\varphi(t-s)|}{1 + |s|} \leq C M(\rho_\alpha\varphi)(t)$$

where C depends only on α .

Proof. Let f be an arbitrary continuous differentiable function. Then the mean value theorem implies

$$\begin{aligned} |f(t)| &\leq \min_{|u-s| \leq 1} |f(u)| + 2 \sup_{|u-s| \leq 1} |Df(u)| \\ &\leq \frac{1}{2} \int_{s-1}^{s+1} |f(u)| du + 2 \sup_{|u-s| \leq 1} |Df(u)| \end{aligned}$$

for all $t, s \in \mathbb{R}$ with $|s - t| \leq 1$. Hence, for $\delta \in (0, 1]$ we deduce

$$|f(t)| \leq \frac{1}{2\delta} \int_{s-\delta}^{s+\delta} |f(u)| du + 2\delta \sup_{|u-s| \leq \delta} |Df(u)| \quad (\text{A.7})$$

for all $t, s \in \mathbb{R}$ with $|s - t| \leq \delta$. Then we apply (A.7) to φ and combine this with (A.4) for ρ_α . We obtain

$$\begin{aligned}
& \rho_\alpha(t-s)|\varphi(t-s)| \\
& \leq \frac{1}{2\delta} \int_{-\delta}^{\delta} \rho_\alpha(t-s-u)|\varphi(t-s-u)| du + 2\delta \sup_{|u| \leq \delta} \rho_\alpha(t-s-u)|D\varphi(t-s-u)| \\
& \leq \frac{1}{2\delta} \int_{|v| \leq 1+|s|} \rho_\alpha(t-v)|\varphi(t-v)| dv + 2\delta \sup_{|u| \leq \delta} \rho_\alpha(t-s-u)|D\varphi(t-s-u)| \\
& \leq \frac{1+|s|}{\delta} M(|\rho_\alpha \varphi|)(t) + 2\delta \sup_{|u| \leq \delta} \rho_\alpha(t-s-u)|D\varphi(t-s-u)|.
\end{aligned}$$

Now, we divide both sides by $1+|s|$, use (A.6) and apply Lemma A.2 to the second summand on the right hand. Then taking the supremum yields

$$\sup_{s \in \mathbb{R}} \frac{\rho_\alpha(t-s)|\varphi(t-s)|}{1+|s|} \leq \frac{1}{\delta} M(|\rho_\alpha \varphi|)(t) + C\delta \sup_{v \in \mathbb{R}} \frac{\rho_\alpha(t-v)|D\varphi(t-v)|}{1+|v|}.$$

Finally, we can choose $\delta < \min\{2/C, 1\}$ which proves the asserted inequality. \square

Proof of Proposition A.1. In a first step, let us prove the right inequality of (A.1) for $\varphi \in S(\mathbb{R})$ satisfying $\text{supp } \varphi \subset [0, 1]$ and an arbitrary $\alpha \in \mathbb{R}$. Again, by the mean value theorem for $\mathcal{F}\varphi$ and property (A.4) we have

$$\rho_\alpha(2\pi k)|\mathcal{F}\varphi(2\pi k)| \leq \rho_\alpha(t)|\mathcal{F}\varphi(t)| + C \sup_{|t-s| \leq 2\pi} \rho_\alpha(s)|D\mathcal{F}\varphi(s)|$$

for all $t \in [2\pi k, 2\pi(k+1)]$. Using $(a+b)^2 \leq 3(a^2 + b^2)$ Lemma A.3 implies

$$\begin{aligned}
& 2\pi \sum_{k \in \mathbb{Z}} \rho_\alpha(2\pi k)^2 |\mathcal{F}\varphi(2\pi k)|^2 \\
& \leq \int_{\mathbb{R}} \rho_\alpha(t)^2 |\mathcal{F}\varphi(t)|^2 dt + C \int_{\mathbb{R}} \sup_{|t-s| \leq 2\pi} \rho_\alpha(s)^2 |D\mathcal{F}\varphi(s)|^2 dt \\
& \leq \|\rho_\alpha \mathcal{F}\varphi\|_{L_2(\mathbb{R})}^2 + C \int_{\mathbb{R}} [M(|\rho_\alpha \mathcal{F}\varphi|)(t)]^2 dt \stackrel{(A.2)}{\leq} C \|\rho_\alpha \mathcal{F}\varphi\|_{L_2(\mathbb{R})}^2.
\end{aligned}$$

Next, we consider again $\alpha \in (1/2, 3/2)$ and we use the fact that $D((0, 1))$ is dense in the Bessel potential spaces $H_2^{-\alpha}((0, 1)) := \{f \in D'((0, 1)) : \|\rho_{-\alpha} \mathcal{F}f\|_{L_2(\mathbb{R})}\}$ which is the dual of $H_{2,0}^\alpha((0, 1))$. Then it follows

$$\begin{aligned}
\|\rho_\alpha \mathcal{F}\varphi\|_{L_2(\mathbb{R})} &= \sup_{\substack{\|\rho_{-\alpha} \mathcal{F}\psi\|_{L_2(\mathbb{R})} \leq 1 \\ \psi \in D((0,1))}} \int_0^1 \varphi(t)\psi(t) dt \\
&= \sup_{\substack{\|\rho_{-\alpha} \mathcal{F}\psi\|_{L_2(\mathbb{R})} \leq 1 \\ \psi \in D((0,1))}} \int_0^1 \sum_{k \in \mathbb{Z}} \hat{\varphi}(k) e^{2\pi kt} \sum_{l \in \mathbb{Z}} \hat{\psi}(l) e^{2\pi lt} dt \\
&= \sup_{\substack{\|\rho_{-\alpha} \mathcal{F}\psi\|_{L_2(\mathbb{R})} \leq 1 \\ \psi \in D((0,1))}} \sum_{k \in \mathbb{Z}} \hat{\varphi}(k) \hat{\psi}(-k)
\end{aligned}$$

$$\begin{aligned}
&\leq \sup_{\substack{\|\rho - \alpha \mathcal{F}\psi\|_{L_2(\mathbb{R})} \leq 1 \\ \psi \in D((0,1))}} \|\rho_\alpha(k) \hat{\varphi}(k)\|_{l_2(\mathbb{Z})} \|\rho_{-\alpha}(k) \hat{\psi}(k)\|_{l_2(\mathbb{Z})} \\
&\leq C \|\rho_\alpha(k) \hat{\varphi}(k)\|_{l_2(\mathbb{Z})}
\end{aligned}$$

where the last estimate is an application of the inequality proved first. This finishes the proof. \square

The technical ideas used above work also in a much more general setting. We refer the reader to [ST87] from which we adapted the proofs for our special situation.

Finally, we remark that $D((0,1))$ is dense in $H_{2,0}^\alpha(\mathbb{T})$, $\alpha \in (1/2, 3/2)$, too. Assume that $f \in H_{2,0}^\alpha(\mathbb{T})$, $f \neq 0$, is orthogonal on $D((0,1))$. This happens if and only if the distribution $L_f \in D'(\mathbb{T})$ defined by

$$L_f(\varphi) := \sum_{k \in \mathbb{Z}} (1 + |k|^2)^\alpha \hat{f}(k) \hat{\varphi}(k), \quad \varphi \in D(\mathbb{T})$$

is supported in $\{0\}$. Yet, in this case L_f has to be a finite linear combination of δ_0 and derivatives of δ_0 . Suppose $L_f := \sum_0^N a_j D^j \delta_0$, $a_N \neq 0$, then it follows that

$$\hat{f}(k) = \frac{1}{(1 + |k|^2)^\alpha} \sum_{j=0}^N a_j k^j.$$

On the other hand side, $f \in H_{2,0}^\alpha(\mathbb{T})$ implies

$$\sum_{k \in \mathbb{Z}} \frac{|a_n|^2 k^{2N}}{(1 + |k|^2)^\alpha} \leq \sum_{k \in \mathbb{Z}} (1 + |k|^2)^\alpha |\hat{f}(k)|^2 < \infty$$

and, consequently, $2N < 2\alpha - 1 < 2$. Hence, we have $\hat{f}(k) = a_0(1 + |k|^2)^{-\alpha}$. Then the condition $0 = f(0) = \sum a_0(1 + |k|^2)^{-\alpha}$ forces a_0 to be zero, which contradicts the assumption $f \neq 0$. Thus, the second statement of Proposition 1.2 follows from Proposition A.1.

References

- [And55] T. W. Anderson. The integral of symmetric unimodular functions over a symmetric set and some probability inequalities. *Proc. Amer. Math. Soc.*, 6:170–176, 1955.
- [Bas88] R. F. Bass. Probability estimates for multiparameter Brownian processes. *Ann. Probab.*, 16(1):251–264, 1988.
- [BM91] A. A. Borokov and A. A. Mogulskii. On probabilities of small deviations for stochastic processes. *Siberian Advances in Math.*, 1:39–63, 1991.
- [Car81] B. Carl. Entropy numbers, s - numbers and eigenvalue problems. *J. Funct. Anal.*, 41:290–306, 1981.
- [CKP] B. Carl, I. Kyrezi, and A. Pajor. Metric entropy of convex hulls in Banach spaces. preprint.
- [CKR93] Z. Ciesielski, G. Kerkycharian, and B. Roynette. Quelques espace fonctionnel associé à des processus Gaussiens. *Stud. Math.*, 107(2):171–204, 1993.
- [CS90] B. Carl and I. Stephani. *Entropy, Compactness and the Approximation of Operators*. Cambridge University Press, 1990.
- [Csá82] E. Csáki. On small values of the square integral of a multiparameter Wiener process. In *Proc. of the 3rd Pannonian Symp. on Math. Stat., Hungary, 1982*, Statistics and Probability, pages 19–26, 1982.
- [DKLL98] T. Dunker, T. Kühn, M. Lifshits and W. Linde. Metric entropy of integration operators and small ball probabilities for the brownian sheet. *C. R. Acad. Sci. Paris, Ser. I*, 326:347–352, 1998.
- [DÜ] L. Decreusefond and A. S. Üstünel. Stochastic analysis of the fractional Brownian motion. preprint.
- [Fel64] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 2. John Wiley & Sons, 1964.
- [KL93] J. Kuelbs and W. V. Li. Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.*, 116:133–157, 1993.
- [KS89] B. S. Kashin and A. A. Saakyan. *Orthogonal Series*, volume 75 of *Transl. of math. Monographs*. AMS, 1989.
- [Li92] W. V. Li. Comparison results for the lower tail of Gaussian seminorms. *J. Theor. Probab.*, 5:1–31, 1992.
- [LL] W. V. Li and W. Linde. Approximation, metric entropy and small ball estimates for Gaussian measures. preprint.

- [LT86] M. A. Lifshits and B. S. Tsirelson. Small ball deviations of Gaussian fields. *Theor. Prob. Appl.*, 31:557–558, 1986.
- [LT91] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: Isoperimetry and Processes*. Springer Verlag, Berlin, 1991.
- [MR95] D. Monrad and H. Rootzén. Small values of Gaussian processes and functional laws of iterated logarithm. *Probab. Theory Relat. Fields*, 101:173–192, 1995.
- [Pie80] A. Pietsch. *Operator Ideals*. North Holland, 1980.
- [Pie87] A. Pietsch. *Eigenvalues and s -Numbers*. Cambridge University Press, 1987.
- [Pis89] G. Pisier. *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge University Press, 1989.
- [PTJ86] A. Pajor and N. Tomczak-Jaegermann. Subspaces of small codimension of finite dimensional Banach spaces. *Proc. Amer. Math. Soc.*, 97:637–642, 1986.
- [Rév81] P. Révész. How small are the increments of a Wiener sheet. In *The 1st Pannonian Symp. on Math. Stat., Bad Tatzmannsdorf, 1979*, volume 8 of *Lecture Notes in Statistics*, pages 206–219, 1981.
- [Sin94] P. Singer. An integrated fractional Fourier transform. *J. Comp. Appl. Math.*, 54:221–237, 1994.
- [ST87] H.-J. Schmeisser and H. Triebel. *Topics in Fourier Analysis and Function Spaces*. Akademische Verlagsgesellschaft Geest & Portig KG, 1987.
- [Ste70] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.
- [Sto95] W. Stolz. *Mesures Gaussiennes de Petites Boules et Petites Déviations*. PhD thesis, Université Paul Sabatier Toulouse III, 1995.
- [Sto96] W. Stolz. Some small ball probabilities for Gaussian processes under non-uniform norms. *J. Theor. Probab.*, 9:613–630, 1996.
- [Tal94] M. Talagrand. The small ball problem for the Brownian sheet. *Ann. Probab.*, 22(3):1331–1354, 1994.
- [Tem90] V. N. Temlyakov. Estimates of asymptotic characteristics of classes of functions with bounded mixed derivative and difference. *Proc. Steklov Inst. Math.*, 189:161–197, 1990.
- [Tem93] V. N. Temlyakov. *Approximation of Periodic Functions*. Nova Science, 1993.

-
- [Tem95] V. N. Temlyakov. An inequality for trigonometric polynomials and its application for estimating the entropy numbers. *J. Complexity*, 11:293–307, 1995.
- [Tri83] H. Triebel. *Theory of Function Spaces*. Birkhäuser, 1983.

Corrections for the PhD thesis: Small Ball Estimates for the Fractional Brownian Sheet

Thomas Dunker

The estimate in the second formula after formula (2.2)

$$(\mathbb{E} |(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|^2)^{1/2} \leq C \sqrt{\#\{m : |m| = n\}} 2^{-\gamma n/2}$$

is certainly true provided the normal random variables $\Delta^{m,i}(\mathbb{G}_d)$ which form

$$(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x) = \sum_{|m|=n} \Delta^{m,i}(\mathbb{G}_d) u_{m,i}(x) \quad \text{with} \quad i_j = \lfloor x_j 2^{m_j} \rfloor$$

are independent. Yet, this assumption does not hold in general. Adding a condition on the regularity of the covariance of the process \mathbb{G}_d we can close this gap. Keeping the notation of the thesis this can be formulated as follows

Proposition 1 *We assume that σ fulfills $(\sigma 1)$ and the condition*

$$(\sigma^*) \quad |\sigma^2(s, t+h) - \sigma^2(s, t)| < Ch|s-t|^{\gamma-1} \text{ for some constant } C > 0.$$

Then it holds

$$(\mathbb{E} |(\mathcal{S}^n - \mathcal{S}^{n-1})\mathbb{G}_d(x)|^2)^{1/2} \leq C \sqrt{\#\{m : |m| = n\}} 2^{-\gamma n/2}.$$

In order to prove this proposition we start with the following lemma

Lemma 2 *Assuming $(\sigma 1)$ and (σ^*) we have*

$$\sum_{|m|=n} \sum_{|m'|=n} |\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_d) \Delta^{m',i(m')}(\mathbb{G}_d)| \leq C \#\{m : |m| = n\} 2^{(1-\gamma)n},$$

where the index $i_j(m) = \lfloor x_j 2^{m_j} \rfloor$ for an arbitrary but fixed $x \in [0, 1]^d$.

Proof. Lemma 1.5 [thesis] showed that we can consider each coordinate of the unit cube separately and that we can combine the results. Any two intervals $[i_j 2^{-m_j}, (i_j + 1) 2^{-m_j})$ and $[i'_j 2^{-m'_j}, (i'_j + 1) 2^{-m'_j})$ which can appear in our situation contain the common fixed point x_j . Consequently, the smaller one is always contained in the bigger one. Let $m \geq m' \geq 0$ then (1.12) – (1.14) [thesis] in combination with (σ^*) yields

$$|\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_1) \Delta^{m',i(m')}(\mathbb{G}_1)| \leq C 2^{(m+m')/2} 2^{-m} 2^{-m'(\gamma-1)}.$$

This is also true when we allow $-1 \leq m' \leq m$ and replace m and m' by $\max(0, m)$ and $\max(0, m')$, respectively. Regrouping the exponents we have

$$|\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_1) \Delta^{m',i(m')}(\mathbb{G}_1)| \leq C 2^{(1-\gamma)(m+m')/2} 2^{-(1-\gamma/2)(m-m')}$$

Observe that $1 - \gamma/2 > 0$ for $\gamma \in (0, 2)$. Applying (1.15) [thesis] we obtain for $|m| = |m'| = n$

$$|\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_d) \Delta^{m',i(m')}(\mathbb{G}_d)| \leq C 2^{(1-\gamma)n} q^{\sum_{j=1}^d |m_j - m'_j|},$$

where $q = 2^{-(1-\gamma/2)} < 1$. Now we fix m with $|m| = n$, let $m' \in M_n$ vary and conclude

$$\begin{aligned} \sum_{|m'|=n} |\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_d) \Delta^{m',i(m')}(\mathbb{G}_d)| &\leq C 2^{(1-\gamma)n} \sum_{|m'|=n} q^{\sum_{j=1}^d |m_j - m'_j|} \\ &\leq C 2^{(1-\gamma)n} \sum_{k=1}^n \sum_{\sum_{j=1}^d |m_j - m'_j| = 2k} q^{2k} \\ &\leq C 2^{(1-\gamma)n} \sum_{k=1}^n (2k)^{d-1} q^{2k} \\ &\leq C 2^{(1-\gamma)n}. \end{aligned}$$

This implies

$$\begin{aligned} \sum_{|m|=n} \sum_{|m'|=n} |\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_d) \Delta^{m',i(m')}(\mathbb{G}_d)| &\leq \sum_{|m|=n} C 2^{(1-\gamma)n} \\ &\leq \#\{m : |m| = n\} C 2^{(1-\gamma)n}, \end{aligned}$$

as asserted. \square

Proof of Proposition 1. Using $\|u_{m,i}\|_{C([0,1]^d)} \leq 2^{-|m|/2}$ and Lemma 2 we have

$$\begin{aligned} \mathbb{E} \left(\sum_{|m|=n} \Delta^{m,i(m)}(\mathbb{G}_d) u_{m,i}(x) \right)^2 &\leq \sum_{|m|=n} \sum_{|m'|=n} |\mathbb{E} \Delta^{m,i(m)}(\mathbb{G}_d) \Delta^{m',i(m')}(\mathbb{G}_d)| 2^{-n} \\ &\leq C \#\{m : |m| = n\} 2^{(1-\gamma)n} 2^{-n} \\ &= C \#\{m : |m| = n\} 2^{-\gamma n}, \end{aligned}$$

which completes the proof. \square

As consequence of this error we have to add the condition (σ^*) to Proposition 2.2, 2.3, Theorem 2.5 and Corollary 2.6 [thesis]. Of course, it would be desirable to find a proof based on a milder condition.

In addition let us correct some typing errors which have been discovered recently. On page 36 in Theorem 3.6 the exponent of ε should be -2 instead of 2 and on page 43 in the second line it must be $2^8(n+1)^{-1} \dots$ and in the fourth line $C 2^{-8}(n+1) \dots$